January, 1966

A GENERALIZATION OF THE GROUP TESTING PROBLEM*

by Satindar Kumar

Technical Report No. 70

University of Minnesota
Minneapolis, Minnesota

*Research supported by National Science Foundation Grant No. NSF-GP-3813.

TABLE OF CONTENTS

In Chapter 1, the 3-category problem with known $\vec{q} = (q_1, q_2, q_3)$ is considered. In section 1.1 the formulation of the problem is given. In section 1.2 to 1.4 a non-mixing procedure R_1 is defined and the properties of R_1 which are concerned with the size of the next test group are investigated. In section 1.5 another procedure R_2 is defined in which the mixing of units from two particular sets is allowed.

Chapter 2 deals with the properties of the optimal procedure R_0 and the lower bounds for any group-testing procedure. In section 2.1 we let $q_2^* = (1/2)[q_2 - 1 + \sqrt{5q_2^2 - 6q_2 + 5}]$ and define $E(T|\vec{q}, N, R_0) = E(T)$ as the expected number of tests to be performed under the optimal procedure R_0 for a starting value of N units and for given $\vec{q} = (q_1, q_2, q_3)$. It is shown that (i) for $q_1 < q_2^*$, $E(T) = N$ and the units are tested one at a time, (ii) for $q_1 > q_2^*$, $E(T) < N$ and (iii) for $q_1 = q_1^*$, $E(T) = N$ and the optimal procedure is not unique. In section 2.2 it is shown that the procedure R_1 is equivalent to the procedure R_0 whenever $q_1 \leq q_2^*$. In section 2.3 we show that the procedure R_1 is better than the procedure R_3 (a modification of R_1). Sections 2.4 and 2.5, respectively, deal with the lower bounds for any group-testing procedure using methods of information theory and coding theory.

In chapter 3, the 3-category problem is extended to the case in which the probabilities q_1, q_2 and q_3 are unknown. In section 3.2 we define a procedure R_1^* --a modification of R_1. It is suggested under R_1^* that a maximum likelihood estimate of q_1, q_2, and q_3 be formed after each test and the procedure R_1 be used assuming that the most recent estimates are the true values. A Bayes procedure $R^{(1)}$ is defined in section 3.3 using a known prior $\lambda(q_1, q_2)$ on the unit square and a useful property of the procedure $R^{(1)}$ is proved in section 3.4.

Chapter 4 deals with the k category problem. A procedure R_1, similar to the procedure in section 1.2, is defined. A property concerning the size of the next test group is proved in section 4.3. In section 4.4 it is shown that for $k \geq 4$ and $N \geq 2$ the optimal procedure R_0 has the following properties: (i) if

$\sum_{j=k-1}^{k} q_j q_{[j]} + \sum_{j=2}^{k-2} q_j(q_1 + q_j) > q_1^2$, then the units are tested one at a time, (ii) if $\sum_{j=2}^{k} q_j q_{[j]} < q_1^2$, the expected number of tests is less than N; here $q_{[j]} = \sum_{i=1}^{j} q_i$. It is also proved that for $q_1 < \frac{1}{2}$ and any k, the optimal procedure tests one unit at a time.

In Chapter 5 we consider modified problems related to the 3-category problem. In section 5.1 we define a procedure R_{11} for the classification of two distinguishable type of units which can be put in the same test group. In section 5.2 we suppose that there are two experimenters working on a single set of N units simultaneously carrying out group tests and cooperating in such a way as to minimize the expected number of tests required to classify all the N units. A procedure R_{12} for this problem is defined.

Returning to the 3-category problem with known $\vec{q} = \vec{q}_0 = (.90,.05,.05)$, Table I gives the values of the next test group under procedure R_1 for the various situations arising in the classification of $N(\leq 8)$ units. In Table II for the same $\vec{q}_0$ we compare the expected number of tests under procedure R_1 with the information-theory lower bounds for any procedure.

Chapter I

The 3-Category Problem with Nested Dominance

1.1 Formulation of the problem

A finite number N of units are to be classified into one of the three disjoint categories. The three categories are labeled as good, mediocre and defective. A group test is a simultaneous test on x units $(1 \leq x \leq N)$ with one of the following three possible outcomes: (i) all the x units are good, (ii) among the x units at least one is mediocre and none are defective, (iii) at least one of the x units is defective. The term "nested dominance" is used, roughly speaking, in the sense that the mediocre units dominate the good units and the defective units dominate both the good and the mediocre units, as far as the sample outcomes are concerned. For sample outcome (ii) and $x \geq 2$ then we do not know which ones or how many units are mediocre and similarly for sample outcome (iii) and $x \geq 2$. The problem is to define a simple and efficient procedure (or an optimal procedure) for classifying all the N units. Each unit is assumed to represent an independent observation from a trinomial population with known a priori probabilities q_1, q_2 and q_3 (with $q_i \geq 0$ and $q_1 + q_2 + q_3 = 1$) of being good, mediocre or defective respectively.

A procedure R_1 which describes a mode of action for any given value of $\vec{q} = (q_1, q_2, q_3)$ is proposed in section 1.2 and some of its properties are studied in section 1.4. Under the procedure R_1, at any stage of the experiment, the experimenter separates the unclassified units into at most four sets and the units within each of these four sets need not be distinguishable. Another procedure R_2, where the identification of the units in the same group-test is sometimes required, is proposed in section 1.5.

1.2 The Procedure R_1

The procedure R_1 is defined by a number of recursion formulae and boundary conditions. Before writing the formulae for R_1, we shall need some definitions and results. A set of units will be called a defective set if it is known to contain at least one defective unit.

For a set of size m, the conditional probability that Z_1, the number of defective units present, equals z given that the set is defective is

$$(1.2.1) \qquad P\{Z_1 = z | Z_1 \geq 1\} = \frac{\binom{m}{z} q_3^z \, q_{[2]}^{m-z}}{1-q_{[2]}^m} \qquad (z = 1,2,\ldots,m)$$

where $q_{[2]} = q_1 + q_2$.

Let x denote the size of a proper (i.e., non-trivial and non-empty) subset randomly chosen from the defective set of size m and let Z_2 denote the number of defective units present in this subset. Then the probability that it contains at least one defective unit is

$$(1.2.2) \qquad P\{Z_2 \geq 1 | Z_1 \geq 1\} = \sum_{z=1}^{m} \sum_{y=1}^{z} \frac{\binom{z}{y}\binom{m-z}{x-y}}{\binom{m}{x}} \frac{\binom{m}{z} q_3^z \, q_{[2]}^{m-z}}{1-q_{[2]}^m} = \frac{1-q_{[2]}^x}{1-q_{[2]}^m}$$

where we use the hypergeometric identity and we define $\binom{z}{y} = 0$ if $y > z$ or $z < 0$.

Let Y_1 be the chance variable representing the number of mediocre units present in the defective set and Y_2 be the chance variable representing the number of mediocre units present in the proper subset of size x randomly chosen from the defective set of size m. Then, for a defective set of size m, the conditional probability that $Y_1 + Z_1$, the number of mediocre plus the number of defective units present, equals a is

$$(1.2.3) \qquad P\{Y_1 + Z_1 = a | Z_1 \geq 1\} = \frac{\sum_{z=1}^{a} \frac{m!}{z!(a-z)!(m-a)!} q_1^{m-a} q_2^{a-z} q_3^z}{1-q_{[2]}^m}$$

$$= \frac{\binom{m}{a}\left[(q_2 + q_3)^a - q_2^a\right] q_1^{m-a}}{1-q_{[2]}^m} (a = 1,2,\ldots,m) .$$

Hence the probability that a randomly chosen proper subset of size x from the defective set of size m, contains all good units is

$$(1.2.4)\quad P\{Y_2 + Z_2 = 0 | Z_1 \geq 1\} = \sum_{r=1}^{m-x} \frac{\binom{m-r}{x}}{\binom{m}{x}} \frac{\binom{m}{r}[(q_2+q_3)^r - q_2^r]q_1^{m-r}}{1-q_{[2]}^m}$$

$$= q_1^x \sum_{r=1}^{m-x} \frac{\binom{m-x}{r}[(q_2+q_3)^r - q_2^r]q_1^{m-r-x}}{1-q_{[2]}^m}$$

$$= \frac{q_1^x(1-q_{[2]}^{m-x})}{1-q_{[2]}^m} .$$

Thus the probability that a proper subset of size x randomly chosen from a defective set of size m contains at least one mediocre unit and no defective unit (taking the complement of the sum of the results in (1.2.2) and (1.2.4)) is

$$(1.2.5)\quad P\{Y_2 \geq 1,\ Z_2 = 0 | Z_1 \geq 1\} = 1 - \frac{1-q_{[2]}^x}{1-q_{[2]}^m} - \frac{q_1^x(1-q_{[2]}^{m-x})}{1-q_{[2]}^m} = \frac{(1-q_{[2]}^{m-x})(q_{[2]}^x - q_1^x)}{1-q_{[2]}^m} .$$

A set of units will be called a mediocre set if it is known that it contains at least one mediocre unit and no defective unit; let W_1 denote the (random) number of mediocre units it contains. Let x denote the size of a proper subset randomly chosen from a mediocre set of size m and let W_2 denote the number of mediocre units in the subset. Then the probability that the subset has only good units is given by

$$(1.2.6)\qquad P\{W_2 = 0 | W_1 \geq 1\} = \frac{q^x(1-q^{m-x})}{1-q^m}$$

where $q = q_1/q_{[2]}$ and the probability that this proper subset of size x is also a mediocre set is given by

$$(1.2.7)\qquad P\{W_2 \geq 1 | W_1 \geq 1\} = \frac{1-q^x}{1-q^m} .$$

Now we shall introduce two lemmas which are of importance in writing the recursion formulas for the procedure R_1.

Lemma 1: Given a mediocre set of size m and given that a randomly chosen proper subset of size x contains at least one mediocre unit, then the a posteriori distribution of the remaining m-x units is that of a binomial sample with probabilities $q_1/q_{[2]}$ and $q_2/q_{[2]}$ of being good and mediocre, respectively.

Lemma 2: Given a defective set of size m and given that a randomly chosen proper subset of size x contains at least one defective unit, then the a posteriori distribution associated with the remaining m-x units is the same as the original trinomial distribution.

These two lemmas will not be proved here since these are special cases of a more general lemma to be proved later.

A set of units will be called a trinomial set if, given the past history of testing, the a posteriori distribution of this set of units is that of independent trinomial chance variables with common probabilities q_1 of being good, q_2 of being mediocre and q_3 of being defective $(q_1 + q_2 + q_3 = 1)$.

A set of units will be called a conditional binomial set if, given the past history of testing, the a posteriori distribution of this set of units is that of independent binomial chance variables with common probabilities $q = q_1/q_{[2]}$ of being good and $q_2/q_{[2]}$ of being mediocre.

The procedure R_1 requires that at every stage the unclassified units be separated into at most four sets, namely, the trinomial set, the defective set, the mediocre set and the conditional binomial set. At some stages, some of these may be empty. At the outset all sets, except the trinomial set of size N, are empty and at the end all of these four sets are empty.

Let $G_1(n_1;m_2,n_2;m_3,n_3;\vec{q}) = G_1(m_2,n_2;m_3,n_3)$ denote the expected number of group-tests remaining to be performed if the procedure R_1 is used and if, presently, the number of classified units is n_1, the mediocre set is of size m_2, the conditional binomial set is of size $n_2 - m_2$, the defective set is of size m_3 and the trinomial set is of size $n_3 - m_3 = N-n_1-n_2-m_3$; the a priori probability of a unit being good, mediocre and defective are known constants q_1, q_2 and $q_3 = 1-q_1-q_2$, respectively and we are using $\vec{q}$ to denote (q_1, q_2, q_3). For the special case when $m_2 = m_3 = 0$ we use the notation $H_1(n_1;n_2;n_3;\vec{q}) = H_1(n_2;n_3)$. The values of n_i $(i = 1,2,3)$ and $m_i (i = 2,3)$ vary as the procedure R_1 of group testing is carried out; at the start of the experiment $n_1 = m_2 = n_2 = m_3 = 0$ and $n_3 = N$. The situation of unclassified units will be referred as a G-situation or $G(m_2,n_2;m_3,n_3)$ - situation if $\max(m_2,m_3) \geq 2$ and as an H-situation or $H(n_2;n_3)$-situation if $m_2 = m_3 = 0$. The case when $\max(m_2,m_3) = 1$ is excluded in the above definition because the G-situation can be changed into H-situation without any group test (see the boundary conditions below) by classifying the unit in the mediocre set, if any, as mediocre and the unit in the defective set, if any, as defective. To simplify the following recursion formulae we drop the subscript 1 on G's and H's since the procedure R_1 is understood and also write m for m_2, d for m_3, n for n_2 and e for n_3.

Recursion Formulae Defining Procedure R_1. For any H-situation with $e \geq 1$, $n \geq 0$ (and $m = d = 0$) we take a sample of size x from the trinomial set and we then have

$$(1.2.8)\quad H(n;e) = 1 + \min_{1 \leq x \leq e} \{q_1^x H(n;e-x) + (q_{[2]}^x - q_1^x)G(x,n+x;0,e-x) + (1-q_{[2]}^x)G(0,n;x,e)\} .$$

For any H-situation with $e=0$, $n \geq 1$

$$(1.2.9)\qquad H(n;0) = 1 + \min_{1 \leq x \leq n} \{q^x H(n-x;0) + (1-q^x)G(x,n;0,0)\} .$$

With the help of lemma 1, (1.2.6) and (1.2.7), for any G-situation with $m \geq 2$ (and any values of $n \geq m$, $e \geq d \geq 0$) we take a sample of size x from the mediocre set and we then have

$$(1.2.10) \quad G(m,n;d,e) = 1 + \min_{1 \leq x \leq m-1} \left\{ \frac{q^x(1-q^{m-x})}{1-q^m} G(m-x,n-x;d,e) + \frac{1-q^x}{1-q^m} G(x,n;d,e) \right\} .$$

With the help of lemma 2, (1.2.2), (1.2.4) and (1.2.5), for any G-situation with $m = 0$, $d \geq 2$ we take a sample of size x from the defective set and we then have

$$(1.2.11) \quad G(0,n;d,e) = 1 + \min_{1 \leq x \leq d-1} \left\{ \frac{q_1^x(1-q_{[2]}^{d-x})}{1-q_{[2]}^d} G(0,n;d-x,e-x) + \frac{(q_{[2]}^x - q_1^x)(1-q_{[2]}^{d-x})}{1-q_{[2]}^d} G(x,n+x;d-x,e-x) + \frac{1-q_{[2]}^x}{1-q_{[2]}^d} G(0,n;x,e) \right\} .$$

The boundary conditions state that for all $\vec{q} = (q_1,q_2,q_3)$

$$(1.2.12) \qquad H(0;0) = 0 .$$

$$(1.2.13) \qquad G(1,n;d,e) = G(0,n-1;d,e) \quad \text{for} \quad n \geq 1,\ e \geq d \geq 0 .$$

$$(1.2.14) \qquad G(0,n;1,e) = H(n;e-1) \quad \text{for} \quad n \geq 0,\ e \geq 1 .$$

In (1.2.8) to (1.2.11) the expression in the braces is the conditional expected number of additional group-tests required to classify all units under procedure R_1 given the size x of the next group-test. It follows from (1.2.8), (1.2.12), (1.2.13) and (1.2.14) that $H(0;1) = 1$ for all $\vec{q}$.

Remark 1: To justify writing $G(x,n;d,e)$ on the right side of (1.2.10) we make use of lemma 1. If the proper subset of size x randomly chosen from the mediocre set of size m is known to contain at least one mediocre unit, then the a posteriori distribution associated with the remaining m-x units is exactly the same as the distribution associated with m-x independent units in the conditional binomial set. These m-x units are then recombined with n-m units in the conditional binomial set giving a total of n-x conditional binomial units, and this justifies the expression $G(x,n;d,e)$ in (1.2.10).

To justify writing $G(0,n;x,e)$ in (1.2.11) we make use of lemma 2. If the proper subset of size x randomly chosen from the defective set of size d is known to contain at least one defective unit, then the a posteriori distribution associated with the remaining d-x units is exactly the same as the distribution associated with d-x independent units in the trinomial set. These d-x units are then recombined with e-d units in the trinomial set, and this justifies the expression $G(0,n;x,e)$ in (1.2.11).

<u>Remark 2</u>: These four recursion formulae, together with boundary conditions allow one to compute successively for any $\vec{q} = (q_1, q_2, q_3)$ the functions $H(0;1)$, $G(2,2;0,0)$, $G(0,0;2,2)$, $H(0;2)$, $G(2,2;0,1)$, $G(2,3;0,0)$, $G(3,3;0,0)$, $G(0,0;3,3)$, $H(0;3)$,...... to any desired values of m, n, d and e.

<u>Remark 3</u>: The integer x which accomplishes the minimization in (1.2.8), (1.2.9), (1.2.10) and (1.2.11) for each situation characterized by the integers m, n, d and e is particularly important, since this is the size of the next group to be tested according to the procedure R_1. These integers $x = x_H(n;e;\vec{q})$ and $x = x_G(m,n;d,e;\vec{q})$ implicitly define the procedure R_1. An illustration of how the procedure R_1 is to be carried out for some explicit values of N and $\vec{q}$ will be given in the next section.

<u>Remark 4</u>: If $m = d = 0$, $e \geq 1$ then it follows from (1.2.7) that under procedure R_1 a subset of size x with $1 \leq x \leq e$ is taken from the trinomial set <u>without mixing it with units from the conditional binomial set</u>. If $m > 1$, then it follows from (1.2.10) that under procedure R_1 a subset of size x with $1 \leq x < m$ is taken from the mediocre set <u>without mixing it with units from the other sets</u>. If $m = 0$, $d > 1$, then it follows from (1.2.11) that under procedure R_1 a subset of size x with $1 \leq x < d$ is taken from the defective set <u>without mixing it with units from</u>

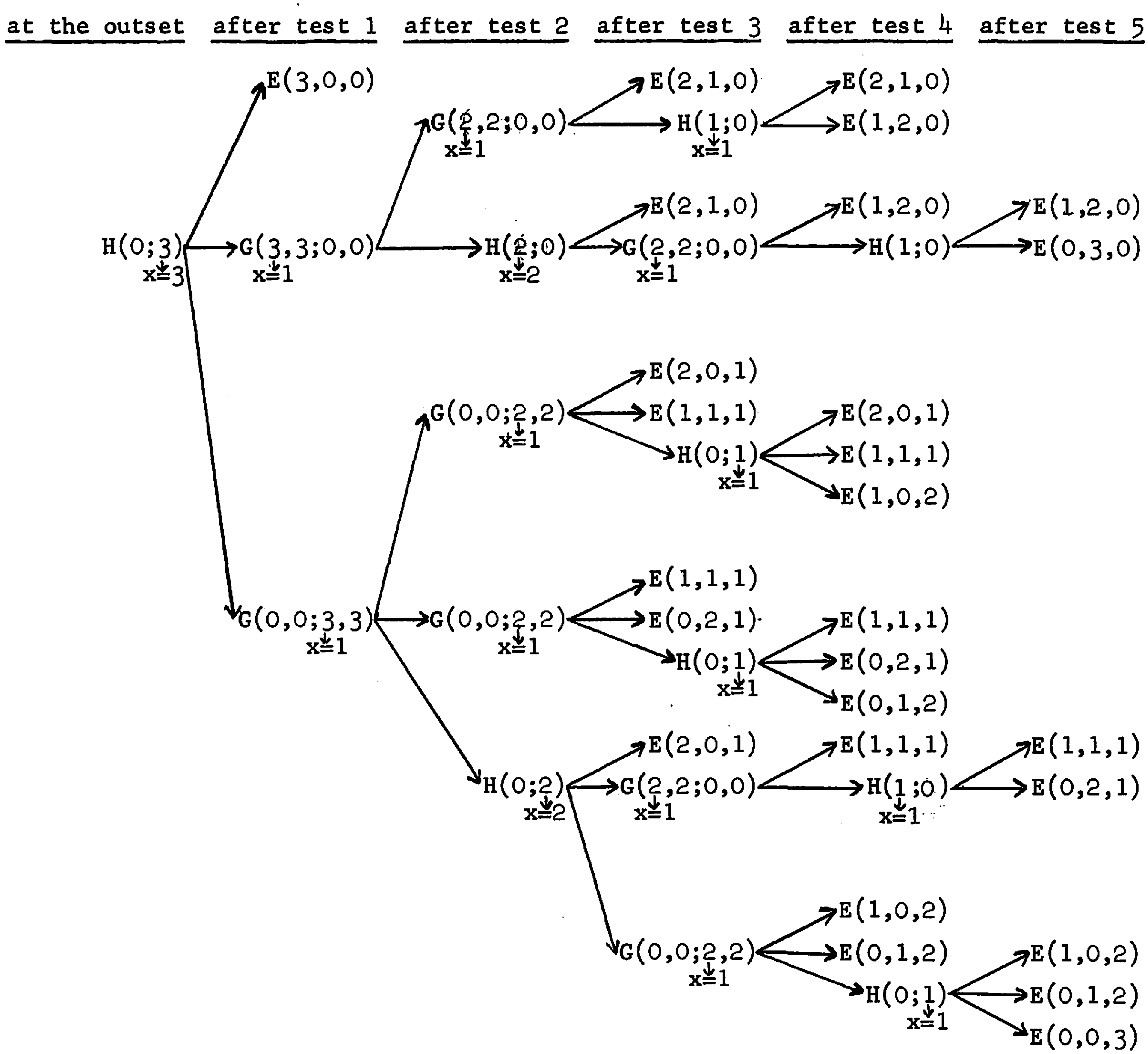
the other sets. It follows from (1.2.7) to (1.2.14) that any lack of optimality for procedure R_1 can only arise from this "no mixing" assumption. Another procedure, which partially drops the "no mixing" assumption at the expense of more complication is introduced in a later section.

1.3. Illustration of the Procedure R_1

Suppose we start with $N = 3$ units and it is given that $q_1 = .90$, $q_2 = .05$ and $q_3 = .05$. The first test-group under procedure R_1 is of size 3. If the test shows all the units to be good, the experiment is over. If these units form a mediocre set (defective set), then accordingly we test a single unit taken randomly from the mediocre set (defective set) under procedure R_1. Similarly we continue along one of the sample paths shown in the figure. The complete tree is shown here for this problem; it consists of at most 5 tests with $3^3 = 27$ end points.

Figure I

Complete tree for procedure R_1
for $q = (.90, .05, .05)$ and $N = 3$



Arrows slanting upwards indicate a good outcome, horizontal arrows indicate a mediocre outcome and arrows slanting downward indicate a defective outcome. $E(i,j,k)$ indicates an outcome with i good units, j mediocre units and k defective units.

1.4 Some Properties of R_1

In this section we consider three properties of the procedure R_1 which are concerned with the size of the next test group; in particular we are interested to determine what this size depends on. These properties are similar to the corresponding properties shown for the analogous procedure for the binomial problem in [5].

Property 1: We state this property as a

Theorem:

For any $G(m,n;d,e)$-situation with $m \geq 2$ and any $\vec{q}$, the size of the next group test under the procedure $R^{(1)}$, defined by (1.2.10) does not depend on n, d or e.

Proof: This proof is a special case obtained by setting $k = 3$ in the proof of the theorem in section 4.3.

Property 2: It will now be shown that in $H(n;e)$-situation with $e \geq 1$, the size x of the next group test taken from the trinomial set of size e may also depend on n.

We need only consider a single numerical example. Let $q_1 = .6$, $q_2 = .1$, and $q_3 = .3$. Suppose we start with the situation $H(n;2)$. Let $H^{(x)}(n;2)$, $(x = 1,2)$ be the expected number of group tests when x is the size of the first group (from the trinomial set of size 2). From (1.2.8) - (1.2.14) we obtain

$$H^{(2)}(n;2) - H^{(1)}(n;2) = q_3 - q_1^2 + q_2\, q_{[2]}\{H(n+1;0) - H(n;0)\} . \tag{1.4.1}$$

If x is independent of n, then (1.4.1) must have the same sign for all n. For $n = 0$ it is easy to see by direct calculation that $H^{(2)}(0;2) = H^{(1)}(0;2) = .01 > 0$, which means that for the $H(0;2)$-situation in this problem, the size of the next group test is one. For $n = 1$ we can show similarly that $H^{(2)}(1;2) - H^{(1)}(1;2) = -.03 < 0$, which means for the $H(1;2)$-situation in this problem, the size of the next group test is two. This establishes the result that in the $H(n;e)$-situation with $e \geq 1$, the size of the next group test may depend on the size n of the conditional binomial set.

Property 3: It will now be shown that in the $G(0,n;d,e)$-situation the size x of the next group test taken from the defective set of size d may also depend on the size n of the conditional binomial set (as well as on the size e-d of the trinomial set). We shall only show the first part, for this we consider the following example with e-d=0. Let $q_1 = .78$, $q_2 = .12$ and $q_3 = .10$. Consider the situations $G(0,0;4,4)$ and $G(0,1;4,4)$. It is easy to show numerically that $x = 3$ cannot be the size of the next group test. Let $G^{(x)}(0,i;4,4)$, $(i = 0,1; x = 1,2)$ be the expected number of tests when x is the size of the next group test (from the defective set of size 4). From (1.2.8)-(1.2.11) we obtain for $i = 0$

$$G^{(1)}(0,0;\ 4,4) - G^{(2)}(0,0;\ 4,4) = \frac{1}{1-q_{[2]}^{4}} \quad (q_{[2]}-q_{[2]}^{4})-(1-q_{[2]}^{2})(q_1q_2+2q_2q_{[2]}+1)$$

This equals $-.014 < 0$ which means that $x = 1$ is for this situation the size of the next group test under procedure R_1. Similarly for $i = 1$

$$G^{(1)}(0,1;\ 4,4) - G^{(2)}(0,1;\ 4,4) = \frac{1}{1-q_{[2]}^{4}} [q_{[2]}-q_{[2]}^{4}- \{q_1q_2q_{[2]}q_3+q_2q_{[2]}(1-q_{[2]}^{2})H(2;0)$$

$$+ 1 - q_{[2]}^{2}+q_1q_2q_3\}] .$$

This equals $.023 > 0$, which means that $x = 2$ is for this situation the size of the next group test under procedure R_1. Hence for the $G(0,n;\ d,e)$ situation, it follows that the size x of the next group test may depend on the size of at least one of the other sets.

Property 4: Using the results of [5], the following results for the situation $G(m,n;\ d,e)$ with $m \geq 2$ are evident.

(i) Under the procedure R_1, with $2 \leq m \leq n$, and $0 \leq \frac{q_1}{q_{[2]}} < .618$,

$$F(m) = \frac{q_1}{q_2} + \frac{q_{[2]}^{m} - q_1^{m-1} q_{[2]} - mq_1^{m}}{q_{[2]}^{m} - q_1^{m}} ,$$

where $F(m)$ is the expected number of group-tests required to "break down" a mediocre set of size m.

(ii) Under the procedure R_1, with $0 \leq \frac{q_1}{q_{[2]}} < .618$, the units from the mediocre sets are tested one at a time until a mediocre unit is found.

(iii) Under the procedure R_1, for the $H(n;0)$-situation with $0 \leq \frac{q_1}{q_{[2]}} < .618$, all the units are tested one at a time.

1.5 Procedure Allowing Mixing of the Trinomial and Conditional Binomial Sets

In this section a new procedure R_2 is introduced where the mixing of the units from the trinomial and conditional binomial sets is allowed whenever we are in an H-situation. For this procedure R_2 three of the four recursion formulas as well as the boundary conditions are the same as for procedure R_1 and will not be repeated here. For the two types of G-situations with $m \geq 2$ or $m = 0$, $d \geq 2$ the procedure R_2 is based on the same recursion formulas, namely, (1.2.10) and (1.2.11), respectively. Under R_2 the recursion formula for the H-situation with $e \geq 1$ is replaced by

$$(1.5.1)\quad H(n;e) = 1+\min\left\{\begin{array}{l}\min\limits_{1\leq x\leq e}\{q_1^x H(n;e-x)+(q_{[2]}^x-q_1^x)G(x,n+x;0,e-x)+(1-q_{[2]}^x)G(0,n;x,e)\} \\ \min\limits_{1\leq y\leq n}\{q_1^e q^y H(n-y;0)+(q_{[2]}^e-q_1^e q^y)G(e+y,n+e;0,0)+(1-q_{[2]}^e)G(0,n;x,e)\}\end{array}\right\}$$

The recursion formula for the H-situation with $e = 0$, $n \geq 1$ is again the same as (1.2.9). The boundary conditions are also the same as in (1.2.12)-(1.2.14).

It follows from the structure (or derivation) of (1.5.1) that under the procedure R_2 for any H-situation with $e \geq 1$ the next test group will be either

(1) a set which is a subset of the trinomial set with no units from the conditional binomial set, or

(2) a set which contains the entire trinomial set and a non-empty subset of the conditional binomial set.

It should be pointed out that in order to apply the procedure R_2 without loss of "information" it is necessary for the H-situation with $e \geq 1$ that the units of the trinomial set must be distinguishable from the units of the conditional binomial set, though the units within each set need not be distinguishable.

Therefore, for any problem in which such an identification of the units is impossible economically or impractical, the procedure R_2 should not be used.

It will now be shown that, in some cases, the expected number of tests under the procedure R_2 is less than the expected number of tests under the procedure R_1. We need consider only a single numerical example. Let $N=3$, $q_1=.65$, $q_2=.25$ and $q_3=.10$. It is easy to show by direct calculation that $x = 2$ is the size of the first group test both under R_1 and R_2. Thus

$$(1.5.2)\quad H_1(0;3) - H_2(0;3) = (q_{[2]}^2 - q_1^2)[G_1(2,2;0,1) - G_2(2,2;0,1)]$$

$$= (q_{[2]}^2 - q_1^2)\left[\frac{q_{[2]}q_2}{q_{[2]}^2 - q_1^2}(H_1(1;1) - H_2(1;1))\right]$$

$$= q_2 q_{[2]}[2 - \{1 + q_3 + (q_1 + q_2 - qq_1)G(2,2;0,0)\}]$$

$$= .046$$

Thus $H_1(0;3) - H_2(0;3) = .046 > 0$, which means that for the $H(0;3)$-situation, with $q_1=.65$, $q_2=.25$ and $q_3=.10$, the expected number of tests under the procedure R_2 is less than the expected number of tests under the procedure R_1.

Chapter 2

Some Properties of the Optimal Procedure R_o and Lower Bounds for any Group-Testing Procedure

2.1 Some Properties of the Optimal Procedure R_o

In this section we discuss some properties of the optimal procedure R_o for the 3-category problem with nested dominance. These properties are concerned with the question of when we should test one-at-a-time and are similar to the corresponding properties shown for the optimal procedure for the binomial problem in [8].

A group-testing procedure is called optimal among all procedures for given N and $\vec{q} = (q_1,q_2,q_3)$ if it minimizes the expected number of tests. Let R_o denote the optimal group-testing procedure and let $E(T|R_o, N, \vec{q}) = E(T)$ denote the expected number of group-tests to be performed if the procedure R_o is used for N units and for given $q = (q_1,q_2,q_3)$. Let q_2^* be defined by

$$(2.1.1) \qquad q_2^* = \tfrac{1}{2}[q_2 - 1 + \sqrt{5q^2 - 6q_2 + 5}] ;$$

we note that for $q_2 = 0$ this reduces to $\frac{1}{2}(\sqrt{5} - 1) = .618$ which plays an important role in the binomial case in [5] and [8].

The optimal procedure R_o for $N = 1$ is, of course, trivial and we now consider R_o for $N \geq 2$.

<u>Theorem</u>:

The optimal procedure R_o has the following properties for $N \geq 2$:

(i) If $q_1 < q_2^*$ then $ET = N$ and the units are tested one at a time.

(ii) If $q_1 > q_2^*$ then $ET < N$.

(iii) If $q_1 = q_2^*$ then $ET = N$ and the optimal procedure is not unique.

Proof: We shall prove (i) by showing that, for $q_1 < q_2^*$, a group-testing procedure cannot be optimal if groups of more than one unit occur at any stage of testing.

A group-testing procedure can be represented by a tree in the following way. The group G_1 on the top represents the group to be tested first. Let G_2 represent the group to be tested next if G_1 turns out to be good, G_3 represent the group to be tested next if G_1 turns out to be mediocre, G_4 represent the group to be tested next if G_1 turns out to be defective. We write G_2, G_3 and G_4 below G_1 and connect it to G_1 by a "path". We can proceed with the representation of the procedure in a similar way.

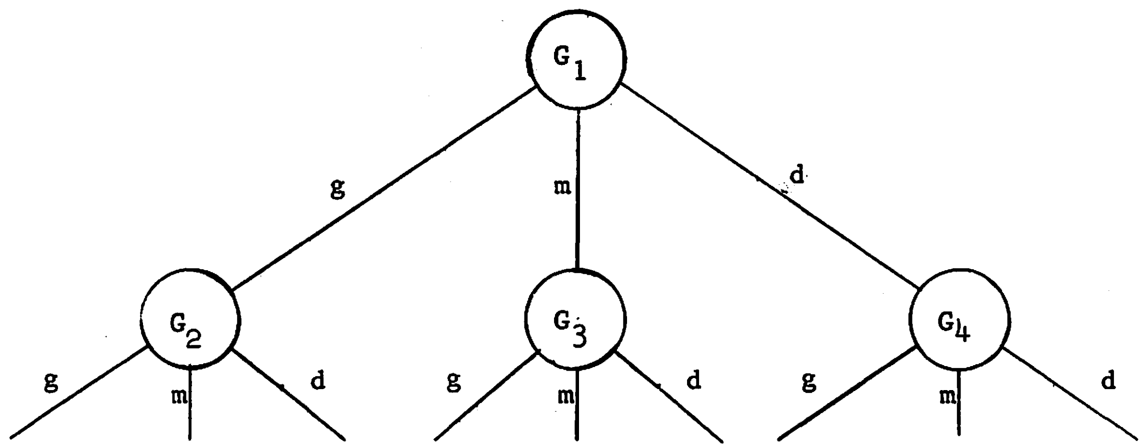


Figure II

Two groups will be called "occuring on the same branch of the tree" if one of them can be reached from the other by descending all the way along one of the connecting paths.

We define a group-testing procedure to be "reasonable" if it has the following properties:

1. A group G will not occur more than once on the same branch of the tree, though it may occur at many places on the same tree. This implies that if a group has already been tested we shall not test the same group again.

2. Let G be the group at any branch point B of the tree. No unit of G has been "previously" classified (i.e., classified at any branch point which has a direct path descending to G).

3. Let G be the group at any branch point B of the tree. There does not exist a

group G' containing G on any of the branches below B.

4. A test will be skipped if the available information at that time enables us to infer the result of the test.

Any group-testing procedure which does not satisfy the above properties can be modified by removing elements from groups and skipping unnecessary tests so as to satisfy these properties. The number of tests needed to classify any sample is definitely not increased by these modifications.

The result (i) of the theorem is proved by considering an arbitrary procedure which satisfies the properties 1 to 4 and modifying it so that the expected number of tests under the new procedure is less than those under the old procedure whenever $q_1 < q_2^*$. We start with a procedure which tests more than one unit at some point and we use the word "plan" to indicate a portion of this procedure. Let B be a branch point

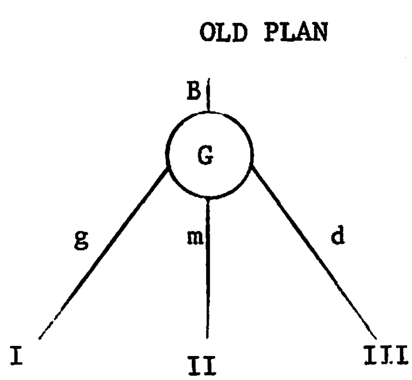
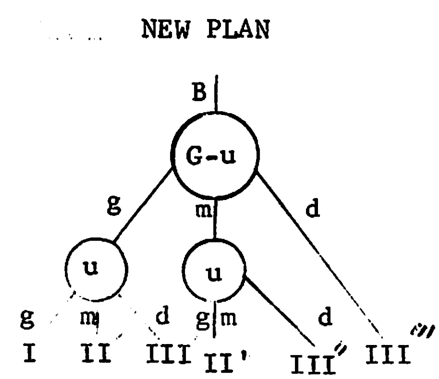


Figure III



on our tree such that the group G to be tested at B has x units where $x \geq 2$ and all the tests below B (if there are any below B) require the units to be tested individually. The branch of the tree is denoted by I, II or III in Fig. III, according as the group test indicates that G is good, mediocre or defective. Let this plan be called the old plan; we now introduce another plan which will be called the new plan. Let u denote a unit of G; u can be any unit of G except that if G is defective then we assume that u is different from the last one to be tested among the units of G under the old plan. Instead of testing G at the branch point B, under the new plan we test G-u. If G-u is found defective we continue as under the old plan where G is found defective. It might be possible to infer the results of some tests and thereby reduce the number of tests by using the additional information available under the new plan. Hence this branch may be different under the new plan and we denote it by III'' instead of III.

If G-u is mediocre, we test the unit u on the next test. If u is found good or mediocre we continue as under the old plan where G is found mediocre. We denote the branch to be followed by II'' instead of II because the availability of the new information again might enable us to infer the results of some tests. If it is found defective we continue as under the old plan where G is found defective. We denote the branch to be followed by III'' because the availability of new information might enable us to infer the results of some test.

If G-u is good, we test u on the next test. If u is good we continue as under the old plan where G is found good. If u is mediocre (or defective) we continue as under the old plan where G is found mediocre (or defective). We denote the branch to be followed by II' (or III') because the availability of new information might enable us to infer the result of some tests. The remainder of the procedure (i.e., everything which is not below B) is left unchanged. The procedure corresponding to old plan (or new plan) is referred to as old procedure (or new procedure).

It is evident from the above construction of the new test plan that, for any sample, the number of tests under the new procedure can at most exceed by one the number of tests for the same sample under the old procedure.

Now we shall show that the only samples for which more tests are needed under the new procedure are those for which B is reached and G is found good under the old plan.

If the branch point B is not reached, then the number of tests for any sample under the old and new procedures are equal since the two procedures are identical elsewhere. Hence in the following discussion we can assume that the branch point B is reached.

If G-u is defective, we follow the same procedure as under the old plan except possibly for skipping a test which may have been necessary under the old plan. Hence, in this case, the number of tests under the new plan is less than or equal to the number of tests under the old plan.

If G–u is mediocre, and u is mediocre or good, then we are following the old plan as in the case where G is mediocre except possibly skipping a test under the new plan which may have been necessary under the old plan. If G–u is mediocre and u is defective, we need $(x + 1)$ tests for classifying all the units of G under the old plan (one test for G and afterward one test for each of the x units of G) whereas under the new plan we shall need either x tests (or $x + 1$ tests) according as it is possible (or not) to infer the result of one test.

If G–u is good and u is mediocre or defective, we need x (or $x + 1$) tests for classifying all the units in G under the old plan according as it is possible (or not) to infer the result of one test whereas we need two tests under the new plan.

If all the units in G are good, we need one test to classify all the units in G under the old plan whereas two tests will be needed under the new plan.

Hence it is established that the only samples for which more tests are needed under the new procedure are those for which B is reached and G is found good under the old procedure.

We discuss the two cases $x = 2$ and x 3 separately:

<u>Case I</u>: <u>$x = 2$</u>

Let $G = (u,v)$. Under the new plan of testing we need two tests to classify the units u and v. Under the old plan of testing the following cases will require three tests to classify the units u and v, i.e., we have a saving of one test for the new plan when

(i) G is defective and the first unit to be tested after G under the old plan along the branch III is defective, and also when

(ii) G is mediocre and the first unit to be tested after G under the old plan along the branch II is mediocre.

In addition we would be using two tests under the new plan when

(iii) G is good whereas we would have used only one test under the old plan (i.e. we have a loss of one test). Combining the results from (i), (ii) and (iii) the expected number of tests saved under the new plan is $(1-q_1-q_2) + q_2(q_1 + q_2) - q_1^2$.

Hence we have a positive saving when

$$(2.1.2) \qquad (1-q_2+q_2^2) - q_1(1-q_2) - q_1^2 \geq 0 \ .$$

Case II: $x \geq 3$

We now show that there is a saving of $x-2$ (or $x-1$) tests when all the $x-1$ units in G except u are good. Under the old plan we would perform $(x-1)$ individual tests to find that all the units in G-u are good and none of these are required under the new plan. Under the old plan a test on the unit u may or may not be necessary but it is necessary under the new plan. Hence in this situation there is a saving of at least $(x-2)$ tests.

Moreover we shall save one test in the following situations:

(1) Let u be a defective unit.

(a) Let b denote the last unit of G-u to be tested under the old plan when G is defective. It will also be the last element of G-u to be tested under the new plan if G-u is defective. If all the elements of G-u, except b, are good we will save one test by using inference in the new plan; this inference is not available to us under the new plan.

(b) Let c be the last element of G-u to be tested when G-u is mediocre. If c is mediocre and all the units in G-u except c are good, then to classify all the units of G under the new plan we need x tests whereas under the old plan of testing we would have needed $(x+1)$ tests to classify the units of G. (The assumption that u is not the last unit tested under the old plan is used here.)

(2) Let u be a mediocre unit.

Let d be the last element of G-u to be tested under the old plan if G is mediocre. It will also be the last unit of G-u to be tested under the new plan if G-u is mediocre. If all the units in G-u, except d, are good we will save one test by using inference in the new plan; this inference is not available to us under the old plan.

Finally if G is good we use two tests under the new plan whereas we would have used only one test under the old plan; hence there is a loss of exactly one test in this case.

Combining all the above results we find that the expected number of tests saved S satisfies the inequality

$$(2.1.3) \qquad S \geq q_1^x \left(\frac{q_2^2 + q_2 q_3 + q_3^2}{q_1^2} + (x-2)\frac{(q_2 + q_3)}{q_1} - 1 \right) .$$

Replacing q_3 by $1-q_1-q_2$ we find that the right hand side of (2.1.3) is positive for all $x \geq 3$ if

$$(2.1.4) \qquad (1-q_2+q_2^2) + q_1(1-q_2) - q_1^2 > 0 .$$

Therefore combining these two cases and noting that the inequalities in (2.1.2) and (2.1.4) are the same, it follows that for all $x (x \geq 2)$ there is a positive saving in the expected number of tests under the new procedure when (2.1.4) holds. Hence the inequality (2.1.4) will be true whenever $q_1 < q_2^*$ where q_2^* is given by (2.1.1).

Furthermore, it is evident that samples with the above mentioned cases will reach the point B. This proves statement (i) of our theorem.

When there are only two units, there are only two different procedures disregarding unreasonable procedures. Under the first procedure we test each unit individually and therefore we need two tests. Under the second procedure, to begin with, we test both units. If both the units are good, we do not need any further test. If this set is mediocre (or defective), we test a single unit. We infer the nature of the second unit, if it is possible to do so, from the result of the test on the first unit; otherwise we test the second unit. The expected number of tests under this second procedure is easily computed to be (e.g., using (1.2.8) - (1.2.14))

$$3 - q_1 - q_2 + q_2(q_1 + q_2) - q_1^2$$

and thus

$$E(T) = \min\{2, 3-q_1-q_2+q_2(q_1+q_2) - q_1^2\} .$$

The second procedure is optimal if

$$(2.1.5) \qquad 1 - q_1 - q_2 + q_2(q_1 + q_2) - q_1^2 < 0$$

or equivalently $q_1 > q_2^*$ where q_2^* is given by (2.1.1). Suppose $N = 2M$ is an even number and (2.1.5) holds. We divide N units into M groups each of size 2 and use the optimal procedure mentioned above for each group of size 2. Under this scheme of testing, the expected number of tests to classify N units is

$$(2.1.6) \qquad M[3 - q_1 - q_2 + q_2(q_1 + q_2) - q_1^2] \ .$$

Now the quantity in square brackets in (2.1.6) is less than 2 and expression (2.1.6) is less than $2M(=N)$. Since the expected number of tests under this procedure is less than N, so it must also be under the optimal procedure. Likewise we can deal with the case $N = 2M + 1$ by dividing these units in $M + 1$ groups, M of which are of size 2 and a group containing a single unit. This proves statement (ii) of our theorem.

It is shown in the proof of the statement (i) of the theorem that, for $x = 2$ the expected number of tests saved under the "new" plan is

$$(2.1.7) \qquad (1-q_1-q_2) + q_2(q_1 + q_2) - q_1^2$$

and for $x \geq 3$, the expected number of tests saved under the new plan is either

$$(2.1.8) \qquad q_1^x\left(\frac{q_2^2 + q_2q_3 + q_3^2}{q_1^2} + (x-1)\frac{(q_2 + q_3)}{q_1} - 1\right)$$

$$= q_1^{x-2}\{(1-q_1-q_2) + q_2(q_1 + q_2) - q_1^2 + (x-2)q_1(q_2 + q_3)\}$$

or

$$(2.1.9) \qquad q_1^x\left(\frac{q_2^2 + q_2q_3 + q_3^2}{q_1^2} + (x-2)\frac{(q_2 + q_3)}{q_1} - 1\right)$$

$$= q_1^{x-2}\{(1-q_1-q_2) + q_2(q_1 + q_2) - q_1^2 + (x-3)q_1(q_2 + q_3)\} \ ,$$

depending on whether or not we can infer the result of one test. For $q_1 = q_2^*$, value of (2.1.7) which equals first three terms in the braces is zero and thus the saving in the expected number of tests is non-negative. The expected number of tests under the new plan is not greater than that under the old plan. Hence introducing a sequence of new plans, each of which tests more units one at a time than the predecessor, it follows that the expected number of tests under the "one-at-a-time" procedure is not greater than the expected number of tests under the procedure arbitrarily chosen at the outset. This proves the statement (iii) of the theorem.

Corollary: For $q_1 < \cdot 6$, the optimal procedure tests one unit at a time.

Proof: Let $q_2 = x$, $q_1 = y$ so that $x + y \leq 1$. The curve corresponding to the equality in (2.1.2), is given by

$$1 - x + x^2 + xy - y - y^2 = 0 \tag{2.1.10}$$

represents a hyperbola, since the discriminant D = 5 is positive. Here we are interested in the upper branch of the hyperbola as shown in Fig.IV. The points of intersection of this branch with the lines $x = o$ and $x + y = 1$ are $(o, \frac{\sqrt{5}-1}{2})$ and $(\frac{3-\sqrt{5}}{2}, \frac{\sqrt{5}-1}{2})$ respectively.

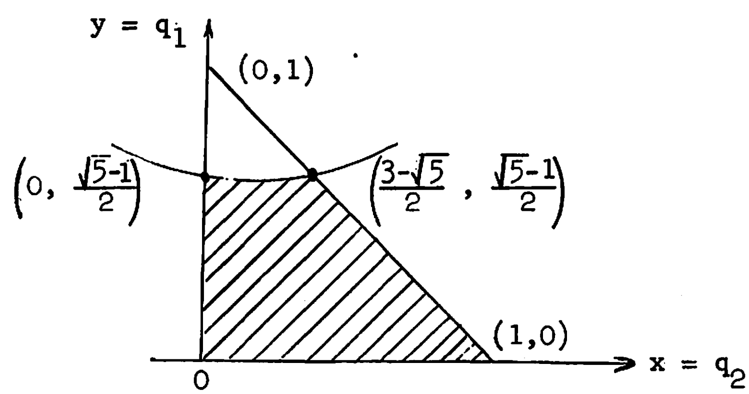


Figure IV

Whenever the point $(q_2, q_1) = (x,y)$ lies in the shaded area, the inequality (2.1.4) is satisfied and the optimal procedure is to test one unit at a time. To obtain the minimum of this branch of the hyperbola, we differentiate (2.1.10) with respect to x, setting $\frac{dy}{dx}$ equal to zero, we obtain

$$2x + y - 1 = 0$$

Substituting $y = 1 - 2x$ in (2.1.10) we obtain the minimum at $y = \frac{3}{5}$ and $x = \frac{1}{5}$. By differentiating (2.1.10) twice with respect to x we find that $\frac{d^2y}{dx^2} = \frac{5}{6} > 0$ at $(\frac{1}{5} , \frac{3}{5})$. Hence the point $(\frac{1}{5} , \frac{3}{5})$ is a minimum for this branch of the hyperbola. Since the minimum is attained at the point $(\frac{1}{5} , \frac{3}{5})$, we are led to the conclusion that when $q_1 < .6$, no matter what the value of q_2 (subject only to $q_1 + q_2 \leq 1$), the optimal procedure for classifying all the units is to test each unit individually. However, for $\frac{3}{5} \leq q_1 < \frac{\sqrt{5}-1}{2}$, we have to check the inequality (2.1.4) and if it holds the optimal procedure is to test each unit individually.

Remarks: For any $\vec{q}$ above the hyperbola, i.e., $q_1 > q_2^*$, we will test more than one unit at a time at some stage since $ET < N$. It is natural to conjecture that for such $\vec{q}$ we would test more than one at a time at the outset but this has not been proved.

2.2 Comparison of the Procedure R_0 with Procedure R_1

In this section we compare the expected number of group tests under the procedure R_0 and R_1 for the classification of N units with known $\vec{q} = (q_1,q_2,q_3)$. Let $E(T|R,N,\vec{q})$ denote the expected number of group tests to be performed if the procedure R is used for N units and for given $\vec{q} = (q_1,q_2,q_3)$. The procedure R is said to be equivalent to the procedure R' for $\vec{q} = (q_1,q_2,q_3)$ if $E(T|R,N,\vec{q}) = E(T|R',N,\vec{q})$ for every N.

We shall use $H(0;N)$ for $E(T|R_1,N,\vec{q})$ where $H(0;N)$ is defined in section 1.2: also q_2^* is the same as defined in (2.1.1).

Theorem: For $\vec{q} = (q_1,q_2,q_3)$ with $q_1 \leq q_2^*$ the procedure R_0 is equivalent to the procedure R_1.

Proof: For $q \leq q_2^*$, $E(T|R_0,N,\vec{q}) = N$ by the theorem in section 2.1 and, since R_0 is the optimal procedure among all procedures, we have

$$E(T|R_0,N,\vec{q}) = N \leq H(0;N). \tag{2.2.1}$$

For any $\vec{q} = (q_1,q_2,q_3)$ and, in particular, for any $\vec{q}$ with $q_1 \leq q_2^*$, we have

(2.2.2)
$$\begin{aligned} H(0;N) &= 1 + \min_{1\leq x\leq N} \{q_1^x H(0;N-1) + (q_{[2]}^x - q_1^x)G(1,1;0,N-1) \\ &\qquad + (1-q_{[2]}^x)G(0,0;1,N)\} \\ &\leq 1 + q_1 H(0;N-1) + (q_{[2]}-q_1)G(1,1;0,N-1) \\ &\qquad + (1-q_{[2]})G(0,0;1,N) \\ &= 1 + H(0;N-1) \quad \text{by using (1.2.13) and (1.2.14).} \end{aligned}$$

For $N = 1$, (2.2.2) gives

$$H(0;1) \leq 1.$$

For $N = 2$, (2.2.2) gives

$$\begin{aligned} H(0;2) &\leq 1 + H(0;1) \\ &\leq 1 + 1 \\ &= 2. \end{aligned}$$

Proceeding in this manner we get

(2.2.3)
$$H(0;N) \leq N.$$

Combining (2.2.1) and (2.2.2) we find for any $\vec{q}$ with $q_1 \leq q_2^*$

$$H(0;N) = E(T|R_0,N,\vec{q}).$$

This proves the theorem.

2.3 Comparison of the Procedure R_1 with Procedure R_3 (Modification of Procedure R_1)

In this section we compare the expected number of group tests under the procedure R_1 and R_3 for the classification of N units with known $\vec{q} = (q_1,q_2,q_3)$.

The procedure R_3 is described as follows: If $N = 2M$ is an even number, we divide the N units into M groups each of size 2 and use the procedure R_1 for each group of size 2. If $N = 2M + 1$ is an odd number, we divide the N units into $M + 1$ groups, M of which are of size 2 and a group containing a single unit and use the procedure R_1 for each group.

The procedure R is said to be better than the procedure R' if $E(T|R,N,\vec{q}) \leq E(T|R',N,\vec{q})$ for every N and every $\vec{q} = (q_1,q_2,q_3)$ and

$E(T|R,N,\vec{q}) < E(T|R',N,\vec{q})$ either for at least one N or for at least one $\vec{q} = (q_1,q_2,q_3)$.

<u>Theorem</u>: The procedure R_1 is better than the procedure R_3.

Proof: For $N = 1$, the two procedures are identical. We now consider R_1 and R_3 for $N \geq 2$. It is shown in the proof of the theorem in section 2.1 that for $q_1 \geq q_2^*$

$$(2.3.1) \qquad E(T|R_3,N,\vec{q}) = \begin{cases} M[3-q_1-q_2+q_2(q_1+q_2)-q_1^2] & \text{if } N = 2M \\ 1+M[3-q_1-q_2+q_2(q_1+q_2)-q_1^2] & \text{if } N = 2M+1. \end{cases}$$

Under the procedure R_1 we obtain from (1.2.8), (1.2.10) and (1.2.11)

$$(2.3.2) \qquad \begin{aligned} H(0;N) &= 1 + \min_{1\leq x\leq N} [q_1^x H(0;N-x)+(q_{[2]}^x-q_1^x)G(x,x;0,N-x) \\ &\qquad +(1-q_{[2]}^x)G(0,0;x,N)] \\ &\leq 1 + q_1^2H(0;N-2)+(q_{[2]}^2-q_1^2)G(2,2;0,N-2) \\ &\qquad +(1-q_{[2]}^2)G(0,0;2,N) \\ &= 1 + q_1^2H(0;N-2)+(q_{[2]}^2-q_1^2)\left\{1+ \frac{q_1q_2}{q_{[2]}^2-q_1^2} H(0;N-2) + \right. \\ &\qquad \left. \frac{q_2q_{[2]}}{q_{[2]}^2-q_1^2} H(1;N-2)\right\} +(1-q_{[2]}^2)\left\{ 1+ \frac{q_{[2]}q_3}{1-q_{[2]}^2} H(0;N-2) + \right. \\ &\qquad \left. \frac{q_3}{1-q_{[2]}^2} H(0;N-1)\right\} \\ &= 2-q_1^2+q_3H(0;N-1)+(q_1+q_2q_3)H(0;N-2)+q_2q_{[2]}H(1;N-2) \end{aligned}$$

It is shown in section 2.2 that for any q_1

$$(2.3.3) \qquad H(0;N-1) \leq 1 + H(0;N-2).$$

We shall prove that

$$(2.3.4) \qquad H(1;N-2) \leq 1 + H(0;N-2).$$

If in the $H(1;N-2)$-situation we continue testing as we would test in the $H(0;N-2)$- situation to reduce the latter to an $H(0;0)$-situation, it is clear from the definition of R_1 that the number of tests under this procedure will be at least $H(1;N-2)$. Since $H(0;N-2)$ tests are required to change the $H(0;N-2)$-situation to an $H(0;0)$-situation, these $H(0;N-2)$ tests will change the $H(1;N-2)$-situation to an $H(1;0)$-situation. Therefore

$$\begin{aligned} H(1;N-2) &\leq H(0;N-2) + H(1;0) \\ &= H(0;N-2) + 1. \end{aligned}$$

Substituting (2.3.3) and (2.3.4) in (2.3.2) we get

$$H(0;N) \leq 3-q_1-q_2+q_2(q_1+q_2)-q_1^2+H(0;N-2). \tag{2.3.5}$$

If $N = 2M$, then by successive application of (2.3.5) we have

$$\begin{aligned} H(0;N) &\leq M[3-q_1-q_2+q_2(q_1+q_2)-q_1^2] \\ &= E(T|R_3,N,\vec{q}) \quad \text{for} \quad q_1 \geq q_2^* . \end{aligned}$$

Similarly for $N = 2M+1$ we get, by using (2.3.5)

$$H(0;N) \leq E(T|R_3,N,\vec{q}) \quad \text{for} \quad q_1 \geq q_2^* .$$

For $q_1 < q_2^*$, R_1 and R_3 are identical for every N and the theorem is proved.

To illustrate the above result we note from Table II that for $N = 4$ and $\vec{q}_0 = (.90,.05,.05)$

$$H(0;4) = 2.034 < 2.576 = E(T|R_3,4,\vec{q}).$$

2.4 Lower Bounds for any Group Testing Procedure from Information Theory

Let $H(N|R)$ be the expected number of group tests needed to classify N units under any arbitrary but fixed procedure R for given $\vec{q} = (q_1,q_2,q_3)$.

Theorem:

$$H(N|R) \geq -N[q_1\log_3 q_1+q_2\log_3 q_2+q_3\log_3 q_3].$$

Proof: The total reduction in entrophy associated with the classification of N units where each unit is assumed to represent an independent observation from a trinomial population with parameter $\vec{q} = (q_1,q_2,q_3)$ is given by

$$(2.4.1) \qquad I = -N\left(\sum_{i=1}^{3} q_i \log_2 q_i\right).$$

The expected number of tests under procedure R, in which the total reduction in entropy is carried out, is $H(N|R)$. The reduction in entropy associated with each test is at most $\log_2 3$; thus we have for any procedure R and any $\vec{q} = (q_1,q_2,q_3)$

$$H(N|R)\log_2 3 \geq -N[q_1\log_2 q_1+q_2\log_2 q_2+q_3\log_2 q_3]$$

or

$$H(N|R) \geq -N[q_1\log_3 q_1+q_2\log_3 q_2+q_3\log_3 q_3].$$

These lower bounds have been calculated for the particular $\vec{q}_0 = (.90,.05,.05)$ and $N \leq 8$ and are given in Table II.

2.5 Lower Bounds for any Group Testing Procedure from Coding Theory

Huffman [4] has given a procedure for the construction of compact codes. Using his results we can obtain a lower bound for any $\vec{q}$ for any group-testing procedure.

Let the set of symbols comprising a given alphabet be called $S = \{s_1,s_2,\ldots,s_q\}$. Then we define a code as a mapping of all possible sequences of symbols of S into sequences of symbols of some other alphabet $X = \{x_1,\ldots,x_r\}$. S is called the source alphabet and X the code alphabet. A compact code for a source S is a code which has the smallest average word length if we encode the symbols from S one at a time.

At the outset there are N units, each of which is good, mediocre or defective. Thus there are 3^N possible states of nature, one of which is true. If we represent each test that gives a good outcome, a mediocre outcome and a defective outcome by the digits zero, one and two respectively, then a procedure is identical with a 3-ary code. Thus a particular set of outcomes (i.e. a particular path in the tree of possible paths) corresponds in a one-to-one manner with a particular "word" of the code. The expected number of tests required is equal to the expected word length of the code.

For example, letting S, T and U denote good, mediocre and defective, respectively, we consider two different codes corresponding to

State of Nature	Probability	Code I	Code II
SS	q_1^2	00	0
ST	q_1q_2	01	10
TS	q_1q_2	10	110
SU	q_1q_3	02	20
US	q_1q_3	20	220
TT	q_2^2	11	111
TU	q_2q_3	12	21
UT	q_2q_3	21	221
UU	q_3^2	22	222

two group-testing procedures for $N = 2$ units. Code I corresponds to the procedure in which each unit is tested individually; code II corresponds to the procedure in which the first test is on both units and subsequent tests are on each unit each.

The expected word length of the code corresponding to any group-testing procedure is clearly greater than or equal to the expected word length of the optimal Huffman code for encoding the 3^N words in the source with known probabilities. Thus the expected word length of the Huffman code is a lower bound on the expected number of tests for any group-testing procedure.

Huffman has given a routine for finding the expected word length of Huffman code. To describe the computation of the expected word length of the Huffman code, let Q_i $(i = 1,2,\ldots,I)$ denote any set of a priori probabilities that sum to one; in our problem of group-testing these a priori probabilities are of a special trinomial structure $q_1^iq_2^jq_3^k$ where i, j, k are non-negative and sum to N. At the 1^{st} step we order the Q_i's, add the three smallest and call the sum S_1; at the 2^{nd} step we reorder the remaining set of 3^N-2 probabilities and again add the three smallest, calling this sum S_2.

This is repeated until the odd number of probabilities remaining reduces to 1. Let S_j denote the sum of three smallest probabilities at the j^{th} step $(j = 1,2,\dots,J)$. It is easy to verify that $J = \frac{1}{2}(3^N-1)$ and $S_J = 1$. Then the Huffman lower bound (HLB) which depends on $\vec{q}$ and N is given by

$$\text{HLB} = \sum_{j=1}^{J} S_j$$

This method is illustrated below by an example with $q_0 = (.90,.05,.05)$ and $N = 2$:

at the outset	after 1st step	after 2nd step	after 3rd step	after 4th step
				1.0000
q_1^2 = .8100	.8100	.8100	.8100	
			.1350	
		.0550	.0550	
q_2q_3 = .0450	.0450	.0450		
q_2q_3 = .0450	.0450	.0450		
q_1q_3 = .0450	.0450	.0450		
q_1q_3 = .0450	.0450			
	.0075			
q_2^2 = .0025	.0025			
q_2q_3 = .0025				
q_2q_3 = .0025				
q_3^2 = .0025				

The value of HLB in the above example is 1.1975 and ILB (Information theory lower bound) is .718 for N = 2. It is easy to see that HLB does not correspond to a group test in this case. The optimal group-test in this case is to start with testing two units at the outset and the expected number of group-tests under the optimal procedure (as well as for procedure R_1) is then easily computed to be 1.2875.

The number of digits (shown below) in a Huffman code word is the number of combinations indicated in the diagram above for the corresponding probability.

State of Nature	Probability	Length of Huffman code word
1	.8100	1
2	.0450	2
3	.0450	2
4	.0450	2
5	.0450	2
6	.0025	2
7	.0025	3
8	.0025	3
9	.0025	3

It is easily observed that the lengths of the Huffman code words (having the same probabilities for the alphabet in the source as the states of nature in our problem) are different from the corresponding word lengths of both code I and code II; the latter corresponds to the optimal group-testing procedure for this problem. Hence the HLB is not attained by a group-testing procedure in this case.

Chapter 3

Maximum Likelihood Solution and Bayes Solution of the Problem

3.1 Introduction

In this chapter we discuss the 3-category problem with nested dominance when $\vec{q} = (q_1, q_2, q_3)$ is unknown. Section 3.2 deals with a maximum likelihood procedure R_1^* which is similar to R_1. Section 3.3 deals with a Bayes procedure $R^{(1)}$ with respect to a known a priori distribution $\lambda(q_1, q_2)$ which is similar to the procedure R_1; section 3.4 deals with a property of the procedure $R^{(1)}$.

3.2 Procedure R_1^* --- A Modification of R_1

It is suggested that a new estimate of $\vec{q} = (q_1, q_2, q_3)$ be formed after each test and then the procedure R_1 be used with the estimated value instead of the true value. To begin with we may use an estimate of $\vec{q} = (q_1, q_2, q_3)$ based on the past experience or arbitrarily start testing one unit at a time. At any stage of the experimentation let s, t and u denote the number of units proven to be good, mediocre and defective, respectively. Also let m, n-m, d and e-d denote the sizes of the mediocre, conditional binomial, defective and trinomial sets, respectively. The notation $q_{[2]} = q_1 + q_2$ and $q = q_1/q_{[2]}$ is the same as in chapter 1. A discussion of the maximum likelihood method of estimation is given below.

Case I: $m \geqq 2$ and $d \geqq 2$

The likelihood function of the observed result is given by

$$L = K\, q_1^s\, q_2^t\, (1-q_{[2]})^u\, q_{[2]}^{n-m}\, (1-q_{[2]}^d)(q_{[2]}^m - q_1^m) \tag{3.2.1}$$

where $K = (s+t+u)!/[s!\ t!\ u!]$. Taking the derivatives of log L with respect to q_1 and q_2 and setting the results equal to zero gives

$$\frac{\partial \log L}{\partial q_1} = \frac{s}{q_1} - \frac{u}{1-q_{[2]}} + \frac{n-m}{q_{[2]}} - \frac{dq_{[2]}^{d-1}}{1-q_{[2]}^d} + \frac{m(q_{[2]}^{m-1} - q_1^{m-1})}{q_{[2]}^m - q_1^m} \doteq 0, \tag{3.2.2}$$

$$(3.2.3)\qquad \frac{\partial \log L}{\partial q_2} = \frac{t}{q_2} - \frac{u}{1-q_{[2]}} + \frac{n-m}{q_{[2]}} - \frac{dq_{[2]}^{d-1}}{1-q_{[2]}^{d}} + \frac{mq_{[2]}^{m-1}}{q_{[2]}^{m}-q_1^{m}} = 0.$$

The roots $\hat{q}_1$ and $\hat{q}_2$, of these equations, are the maximum likelihood estimates of q_1 and q_2, respectively. Subtracting (3.2.3) from (3.2.2), we now obtain

$$(3.2.4)\qquad \frac{s}{\hat{q}_1} - \frac{t}{\hat{q}_2} + \frac{m\hat{q}_1^{m-1}}{\hat{q}_{[2]}^{m}-\hat{q}_1^{m}} = 0$$

The equations (3.2.2) and (3.2.4) form a system equivalent to the equations (3.2.2) and (3.2.3). Multiplying both sides of (3.2.4) by $\hat{q}_{[2]}$ and replacing $\hat{q}_1/\hat{q}_{[2]}$ by $\hat{q}$, we get

$$\frac{s}{\hat{q}} - \frac{t}{1-\hat{q}} - \frac{m\hat{q}^{m-1}}{1-\hat{q}^{m}} = 0$$

or equivalently

$$(3.2.5)\qquad s - t\sum_{i=1}^{m} \hat{q}^{i} - (s+m)\hat{q}^{m} = 0$$

For positive s, using the Descartes Rule of Signs we see that (3.2.5) has exactly one root in the unit interval and hence $\hat{q}$ is uniquely determined and it is easy to check that $\hat{q} < 1$. Then $\hat{q}_1 = \hat{q}\,\hat{q}_{[2]}$; substituting this fixed value of $\hat{q}_1$ in (3.2.2), we obtain after simplification

$$(3.2.6)\qquad \left(\frac{s}{\hat{q}} + n - m + \frac{m(1-\hat{q}^{m-1})}{1-\hat{q}^{m}}\right) - u\sum_{i=1}^{d} \hat{q}_{[2]}^{i} - \left(\frac{s}{\hat{q}} + n - m + \frac{m(1-\hat{q}^{m-1})}{1-\hat{q}^{m}} + d\right)\hat{q}_{[2]}^{d} = 0;$$

which we use to obtain $\hat{q}_{[2]}$. This equation is quite similar to (3.2.5) and it follows as before that there is exactly one root $\hat{q}_{[2]}$ in the unit interval; thus $\hat{\vec{q}} = (\hat{q}_1, \hat{q}_2, \hat{q}_3)$ is uniquely determined.

If $s = 0$, then $\hat{q} = \hat{q}_1 = 0$ and hence $\hat{q}_{[2]} = \hat{q}_2$ can be found from

$$n - u\sum_{i=1}^{d} \hat{q}_{[2]}^{i} - (n+d)\hat{q}_{[2]}^{d} = 0$$

Case II: m = 0 and d ≧ 2

In this case the discussion is quite similar; $\hat{q}$ and $\hat{q}_{[2]}$ are given by the equations

(3.2.7) $$\frac{s}{\hat{q}_1} - \frac{t}{\hat{q}_2} = 0$$

(3.2.8) $$\left(\frac{s}{\hat{q}} + n\right) - u \sum_{i=1}^{d} \hat{q}_{[2]}^{i} - \left(\frac{s}{\hat{q}} + n + d\right) \hat{q}_{[2]}^{d} = 0$$

where again $\hat{q} = \hat{q}_1/\hat{q}_{[2]}$ is obtained directly from (3.2.7). As in case I we conclude that $\hat{\vec{q}} = (\hat{q}_1, \hat{q}_2, \hat{q}_3)$ is uniquely determined.

Case III: m ≧ 2 and d = 0

In this case $\hat{q}$ and $\hat{q}_{[2]}$ are given by the equations

(3.2.9) $$s - t \sum_{i=1}^{m} \hat{q}^{i} - (s + m) \hat{q}^{m} = 0$$

(3.2.10) $$\left(\frac{s}{\hat{q}} + \frac{m(1-\hat{q}^{m-1})}{1-\hat{q}^{m}} + n - m\right) - \left(\frac{s}{\hat{q}} + \frac{m(1-\hat{q}^{m-1})}{1-\hat{q}^{m}} + n - m + u\right) \hat{q}_{[2]} = 0$$

where $\hat{q}_{[2]}$ is directly obtained from (3.2.10). Again, by using Descartes Rule of Signs in (3.2.9) we find that $\hat{\vec{q}} = (\hat{q}_1, \hat{q}_2, \hat{q}_3)$ is uniquely determined.

Case IV: m = 0 and d = 0

In this case $\hat{q}$ and $\hat{q}_{[2]}$ are given by the equations

(3.2.11) $$s - (s + t) \hat{q} = 0,$$

(3.2.12) $$\left(\frac{s}{\hat{q}} + n\right) - \left(\frac{s}{\hat{q}} + n + u\right) \hat{q}_{[2]} = 0.$$

Solving (3.2.11) and (3.2.12) we get

$$\hat{q}_1 = \frac{s(s + t + n)}{(s + t)(s + t + u + n)} \quad \text{and} \quad \hat{q}_2 = \frac{t(s + t + n)}{(s + t)(s + t + u + n)}$$

3.3 Bayes Solution $R^{(1)}$

As in the case of R_1, at any stage of experimentation units not individually classified (under the procedure $R^{(1)}$ defined below) will be kept in at most four distinguishable sets, a trinomial set, a conditional binomial set, a mediocre set and a defective set. Units for any one test will be taken from only one of these four sets; it is for this reason that the procedure $R^{(1)}$ will be sometimes referred to as a non-mixing procedure or as being of R_1-type.

At any stage of the experiment let s, t and u denote the number of units definitely proven to be good, mediocre and defective, respectively; let m, n-m, d and e-d denote the sizes of the current mediocre, conditional binomial, defective and trinomial sets respectively. Here $N = s+t+u+n+e$.

Let $\lambda(q_1,q_2)$ denote a completely known a priori distribution of $\hat{q} = (q_1,q_2)$. Then the a posteriori density element is given for m = 0, 2, 3, ... and d = 0, 2, 3, ... by

$$(3.3.1) \quad d\lambda_{s,t,u}(q_1,q_2;m,n;d) = \frac{q_1^s q_2^t (1-q_{[2]})^u q_{[2]}^{n-m}(q_{[2]}^m - q_1^m h(m))(1-q_{[2]}^d h(m))}{C(s,t,u;\ m,n;d)}$$

where $C(s,t,u;\ m,n;d)$ and $h(z)$ are defined by

$$(3.3.2) \quad C(s,t,u;\ m,n;d) = B(s,t,u,n;m) - B(s,t,u,n+d;m)h(d),$$

$$h(0) = 0,$$

$$h(z) = 1 \qquad \text{for } z \geq 2.$$

Here

$$(3.3.3) \quad B(s,t,u,n;m) = A(s,t,u,n) - A(s+m,t,u,n-m)h(m)$$

where

$$(3.3.4) \qquad A(s,t,u,n) = \int_\Delta q_1^s q_2^t (1-q_{[2]})^u q_{[2]}^n d\lambda(q_1,q_2) ,$$

the range of integration being the triangle Δ given by $q_1 \geq 0$, $q_2 \geq 0$, $q_{[2]} = q_1+q_2 \leq 1$. We also define

$$(3.3.5) \quad C(s,t,u;m,n;1) = B(s,t,u,n;m) - B(s,t,u,n+1;m),$$

$$(3.3.6) \quad B(s,t,u,n;1) = A(s,t,u,n) - A(s+1,t,u,n-1).$$

It is easy to verify the identities (3.3.7)-(3.3.15) for $x \geq 1$.

(3.3.7) $B(s,t,u,n;1) = A(s,t+1,u,n-1)$

(3.3.8) $C(s,t,u;m,n;1) = B(s,t,u+1,n;m)$

(3.3.9) $A(s,t,u,n) = C(s,t,u;0,n;0)$,

(3.3.10) $B(s,t,u,n;m) = C(s,t,u;m,n;0)$,

(3.3.11) $A(s+x,t,u,n) + B(s,t,u,n+x;x) + C(s,t,u;0,n;x) = A(s,t,u,n)$,

(3.3.12) $A(s+x,t,u,n-x) + B(s,t,u,n;x) = A(s,t,u,n)$,

(3.3.13) $C(s+x,t,u;m-x,n-x;d) + C(s,t,u;x,n;d) = C(s,t,u;m,n;d)$,

(3.3.14) $C(s+x,t,u;0,n;d-x) + C(s,t,u;x,n+x;d-x) + C(s,t,u;0,n;x) = C(s,t,u;0,n;d)$,

(3.3.15) $\sum_{j=s}^{s+m-1} C(j,t+1,u;0,n-j+s-1;d) = C(s,t,u;m,n;d)$,

We note that if $\lambda(q_1,q_2)$ is the uniform over the triangle $q_1 \geq 0$, $q_2 \geq 0$, $q_{[2]} \leq 1$, then the integration in (3.3.4) is a Dirichlet density and we easily obtain

(3.3.16) $A(s,t,u,n) = 2 \sum_{k=0}^{n} \binom{n}{k} \frac{(s+k)!(n+t-k)!(u)!}{(s+n+t+u+2)!}$

Equating the coefficient of x^n in $(1-x)^{-(s+1)}(1-x)^{-(t+1)}$ and $(1-x)^{-(s+t+2)}$ we find the right hand side of (3.3.16) is equal to

$$2\,\frac{(s+t+n+1)!s!t!u!}{(s+t+1)!(s+n+t+u+2)!}$$

In (3.3.1) the case $m = d = 0$ will be referred to as an H-situation, the case in which $\max(m,d) \geq 2$ will be referred to as a G-situation. The remaining case in which $\max(m,d) = 1$ is reduced without testing to an H-situation by classifying the unit in the mediocre set, if $m = 1$, as mediocre and/or the unit in the defective set, if $d = 1$, as defective.

The expected number of additional tests required at any stage depends on s,t,u,m,n,d,e and $\vec{q} = (q_1,q_2)$; let $G_{s,t,u}(m,n;d,e) = G^{(1)}_{s,t,u}(m,n;d,e|\vec{q},R^{(1)})$ denote the expected number of group-tests when the procedure $R^{(1)}$ is used; the superscript is dropped since we shall only discuss $R^{(1)}$ in sections 3.3 and 3.4. For the case $m = d = 0$, we shall use $H_{s,t,u}(n;e)$ instead of $G_{s,t,u}(0,n;0,e)$; in all cases we have $s \geq 0$, $t \geq 0$ and $u \geq 0$. We now use these definitions to define the

<u>Bayes Procedure $R^{(1)}$</u> (Non-mixing type).

For any H-situation with $n \geqq 0$ and $e \geqq 1$ we take a sample of size x from the trinomial set so that $1 \leqq x \leqq e$, and

$$(3.3.17)\quad H_{s,t,u}(n;e) = 1+q_1^x H_{s+x,t,u}(n;e-x)+(q_{[2]}^x-q_1^x)G_{s,t,u}(x,x+n;0,e-x)$$
$$+(1-q_{[2]}^x)G_{s,t,u}(0,n;x,e),$$

where x is the integer (with $1 \leqq x \leqq e$) that minimizes the integral on the right hand side of (3.3.17) with respect to the a posteriori density element $d\lambda_{s,t,u}(q_1,q_2;0,n;0)$.

For any H-situation with $n \geqq 1$ and $e = 0$ we take a sample of size x from the conditional binomial set so that $1 \leqq x \leqq n$, and

$$(3.3.18)\quad H_{s,t,u}(n;0) = 1+q^x H_{s+x,t,u}(n-x;0)+(1-q^x)G_{s,t,u}(x,n;0,0),$$

where x is the integer (with $1 \leqq x \leqq n$) that minimizes the integral on the right hand side of (3.3.18) with respect to the a posteriori density element $d\lambda_{s,t,u}(q_1,q_2;0,n;0)$.

For any G-situation with $n \geqq m \geqq 2$ we take a sample of size x from the mediocre set so that $1 \leqq x \leqq m-1$, and

$$(3.3.19)\quad G_{s,t,u}(m,n;d,e) = 1+\left(\frac{q^x-q^m}{1-q^m}\right)G_{s+x,t,u}(m-x,n-x;d,e)+\left(\frac{1-q^x}{1-q^m}\right)G_{s,t,u}(x,n;d,e),$$

where x is the integer (with $1 \leqq x \leqq m-1$) that minimizes the integral on the right hand side of (3.3.19) with respect to the a posteriori density element $d\lambda_{s,t,u}(q_1,q_2;m,n;d)$.

For any G-situation with $m = 0$ and $e \geqq d \geqq 2$ we take a sample of size x from the defective set so that $1 \leqq x \leqq d-1$, and

$$(3.3.20)\quad G_{s,t,u}(0,n;d,e) = 1+\left(\frac{q_1^x(1-q_{[2]}^{d-x})}{1-q_{[2]}^d}\right) G_{s+x,t,u}(0,n;d-x,e-x)$$

$$+\left(\frac{(q_{[2]}^x-q_1^x)(1-q_{[2]}^{d-x})}{1-q_{[2]}^d}\right) G_{s,t,u}(x,n+x;d-x,e-x)$$

$$+\left(\frac{1-q_{[2]}^x}{1-q_{[2]}^d}\right) G_{s,t,u}(0,n;x,e),$$

where x is the integer (with $1 \leqq x \leqq d-1$) that minimizes the integral on the right hand side of (3.3.20) with respect to the a posteriori density element $d\lambda_{s,t,u}(q_1,q_2;0,n;d)$.

The boundary conditions state that for all s,t,u and $\vec{q}$,

$$(3.3.21)\quad H_{s,t,u}(0;0) = 0.$$

$$(3.3.22)\quad G_{s,t,u}(1,n;d,e) = G_{s,t+1,u}(0,n-1;d,e) \text{ for } n \geqq 1,\ e \geqq d \geqq 0.$$

$$(3.3.23)\quad G_{s,t,u}(0,n;1,e) = H_{s,t,u+1}(n;e-1) \text{ for } e \geqq 1,\ n \geqq 0.$$

In each of the above situations the integer x (as mentioned above) is the size of the next group test and the source of these x units has already been described. Thus the knowledge (i.e., a table) of these x-values describes explicitly the procedure $R^{(1)}$.

From (3.3.1) we have

$$(3.3.24)\quad q_1^x d\lambda_{s,t,u}(q_1,q_2;0,n;0) = \frac{A(s+x,t,u,n)}{A(s,t,u,n)}\, d\lambda_{s+x,t,u}(q_1,q_2;0,n;0).$$

$$(3.3.25)\quad (q_{[2]}^x-q_1^x)d\lambda_{s,t,u}(q_1,q_2;0,n;0) = \frac{B(s,t,u,n+x;x)}{A(s,t,u,n)}\, d\lambda_{s,t,u}(q_1,q_2;x,n+x;0).$$

$$(3.3.26)\quad (1-q_{[2]}^x)d\lambda_{s,t,u}(q_1,q_2;0,n;0) = \frac{C(s,t,u;0,n;x)}{A(s,t,u,n)}\, d\lambda_{s,t,u}(q_1,q_2;0,n;x).$$

Now we define the Bayes averages (not depending on (q_1,q_2))

$$(3.3.27)\quad \bar{H}_{s,t,u}(n;e) = \int_\Delta H_{s,t,u}(n;e)d\lambda_{s,t,u}(q_1,q_2;0,n;0).$$

$$(3.3.28)\quad \overline{G}_{s,t,u}(m,n;d,e) = \int_\triangle G_{s,t,u}(m,n;d,e)d\lambda_{s,t,u}(q_1,q_2;m,n;d).$$

Now we shall use (3.3.24)-(3.3.28) to write the recursion formulae (3.3.17)-(3.3.20) in a simpler form without integrals.

Integrating (3.3.17) and (3.3.18) with respect to the a posteriori density element $d\lambda_{s,t,u}(q_1,q_2;0,n;0)$ gives for $e \geqq 1$ and $e = 0$ respectively

$$(3.3.29)\quad \overline{H}_{s,t,u}(n;e) = 1+\min_{1\leq x\leq e} \int_\triangle \{q_1^x H_{s+x,t,u}(n;e-x)+(q_{[2]}^x-q_1^x)G_{s,t,u}(x,x+n;0,e-x)$$
$$+(1-q_{[2]}^x)G_{s,t,u}(0,n;x,e)\}d\lambda_{s,t,u}(q_1,q_2;0,n;0)$$

$$= 1+\min_{1\leq x\leq e} \{\frac{A(s+x,t,u,n)}{A(s,t,u,n)}\overline{H}_{s+x,t,u}(n;e-x) +$$
$$\frac{B(s,t,u,n+x;x)}{A(s,t,u,n)}\overline{G}_{s,t,u}(x,n+x;0,e-x) +$$
$$\frac{C(s,t,u;0,n;x)}{A(s,t,u,n)}\overline{G}_{s,t,u}(0,n;x,e)\},$$

$$(3.3.30)\quad \overline{H}_{s,t,u}(n;0) = 1+\min_{1\leq x\leq n} \int_\triangle \{q^x H_{s+x,t,u}(n-x;0) +$$
$$(1-q^x)G_{s,t,u}(x,n;0,0)\}d\lambda_{s,t,u}(q_1,q_2;0,n;0)$$

$$= 1+\min_{1\leq x\leq n} \{\frac{A(s+x,t,u,n-x)}{A(s,t,u,n)}\overline{H}_{s+x,t,u}(n-x;0) +$$
$$\frac{B(s,t,u,n;x)}{A(s,t,u,n)}\overline{G}_{s,t,u}(x,n;0,0)\}.$$

Integrating both sides of (3.3.19) with respect to the a posteriori density element $d\lambda_{s,t,u}(q_1,q_2;m,n;d)$ we get

$$(3.3.31)\quad \overline{G}_{s,t,u}(m,n;d,e) = 1+\min_{1\leq x\leq m-1} \int_\triangle \{(\frac{q^x-q^m}{1-q^m})G_{s+x,t,u}(m-x,n-x;d,e) +$$
$$(\frac{1-q^x}{1-q^m})G_{s,t,u}(x,n;d,e)\}d\lambda_{s,t,u}(q_1,q_2;m,n;d)$$

$$= 1+\min_{1\leq x\leq m-1} \{\frac{C(s+x,t,u;m-x,n-x;d)}{C(s,t,u;m,n;d)} \overline{G}_{s+x,t,u}(m-x,n-x;d,e)$$

$$+ \frac{C(s,t,u;x,n;d)}{C(s,t,u;m,n;d)} \overline{G}_{s,t,u}(x,n;d,e)\}.$$

Integrating both sides of (3.3.20) with respect to the a posteriori density element $d\lambda_{s,t,u}(q_1,q_2;0,n;d)$, we get

$$(3.3.32)\quad \overline{G}_{s,t,u}(0,n;d,e) = 1+\min_{1\leq x\leq d-1} \int_{\triangle}\{(\frac{q_1^x(1-q_{[2]}^{d-x})}{1-q_{[2]}^d})G_{s+x,t,u}(0,n;d-x,e-x) + (\frac{(q_{[2]}^x-q_1^x)(1-q_{[2]}^{d-x})}{1-q_{[2]}^d})\,G_{s,t,u}(x,n+x;d-x,e-x) + (\frac{1-q_{[2]}^x}{1-q_{[2]}^d})G_{s,t,u}(0,n;x,e)\}d\lambda_{s,t,u}(q_1,q_2;0,n;d)$$

$$= 1+\min_{1\leq x\leq d-1} \{\frac{C(s+x,t,u;0,n;d-x)}{C(s,t,u;0,n;d)} \overline{G}_{s+x,t,u}(0,n;d-x,e-x)$$

$$+ \frac{C(s,t,u;x,n+x;d-x)}{C(s,t,u;0,n;d)} \overline{G}_{s,t,u}(x,n+x;d-x,e-x)$$

$$+ \frac{C(s,t,u;0,n;x)}{C(s,t,u;0,n;d)} \overline{G}_{s,t,u}(0,n;x,e)\}.$$

The boundary conditions state that for all s,t and u

(3.3.33) $\overline{H}_{s,t,u}(0;0) = 0.$

(3.3.34) $\overline{G}_{s,t,u}(1,n;d,e) = \overline{G}_{s,t+1,u}(0,n-1;d,e)$ for $n \geq 1$, $e \geq d \geq 0$.

(3.3.35) $\overline{G}_{s,t,u}(0,n;1,e) = \overline{H}_{s,t,u+1}(n;e-1)$ for $n \geq 0$, $e \geq 1$.

Equations (3.3.29)-(3.3.32) do not involve integrals and it is therefore easier to iterate these equations on a computer for given values of s,t,u,m, n,d,e and the distribution λ.

The functions $H_{s,t,u}(n;e)$, $H_{s,t,u}(n;0)$, $G_{s,t,u}(m,n;d,e)$ and $G_{s,t,u}(0,n;d,e)$ can be expressed as functions of (q_1,q_2). However, these expressions need not be unique, since the Bayes procedure may not be unique. Nevertheless the Bayes averages $\bar{H}_{s,t,u}(n;e)$, $\bar{H}_{s,t,u}(n;o)$, $\bar{G}_{s,t,u}(m,n;d,e)$ and $\bar{G}_{s,t,u}(0,n;d,e)$ are unique.

3.4 A Property of the Procedure $R^{(1)}$

In this section we consider a property of the procedure $R^{(1)}$ which is concerned with the size of the next group-test in a G-situation with $m \geq 2$. This property is similar to a result shown for the analogous procedure in the binomial problem[3] We shall show that for the G-situation with $m \geq 2$, the size of the next test group, defined (3.3.19), does not depend on e. For this situation the immediate objective under procedure $R^{(1)}$ is to break down the mediocre set of size $m \geq 2$ until a single unit is proven to be mediocre and is removed (from the unclassified units). Instead of randomizing the units in this mediocre set each time before a test group is selected, it is assumed without any loss of generality that the order is randomized only once at the outset; then units removed for testing will be taken in that order. If the ith unit is the first mediocre unit, then the breaking down of the mediocre set will lead to a situation in which the mediocre set will be empty, the size of the conditional binomial set will be increased by m-i and the number of units classified as good and mediocre will be increased by i-1 and 1, respectively. These properties are a consequence of the way in which $R^{(1)}$ is defined.

Let $F_{s,t,u}(m,n;d) = F^{(1)}_{s,t,u}(m|\,(q_1,q_2),R^{(1)};n;d,e)$ be defined as the expected number of group-tests required to break down a mediocre set of size m and reach (for the first time) a situation in which the mediocre set is empty, when $\lambda(q_1,q_2)$ is given and the procedure $R^{(1)}$ is used. It is clear from the sub procedure described in the previous paragraph that for any s,t,u,m,n,d the value of $F_{s,t,u}(m,n;d)$ will not depend on e. Then we can write

$$(3.4.1)\quad G_{s,t,u}(m,n;d,e) = F_{s,t,u}(m,n;d)+\left(\frac{1-q}{1-q^m}\right)\sum_{i=1}^{m} q^{i-1}G_{s+i-1,t+1,u}(0,n-i;d,e).$$

As in (3.2.27) and (3.2.28) we define

$$(3.4.2)\quad \bar{F}_{s,t,u}(m,n;d) = \int_{\Delta} F_{s,t,u}(m,n;d)d\lambda_{s,t,u}(q_1,q_2;m,n;d).$$

Integrating both sides of (3.4.1) with respect to the a posteriori density $d\lambda_{s,t,u}(q_1,q_2;m,n;d)$, using (3.4.2) and letting $s+i-1 = j$, we obtain

$$(3.4.3)\quad \overline{G}_{s,t,u}(m,n;d,e) = \overline{F}_{s,t,u}(m,n;d)+\sum_{j=s}^{s+m-1} \frac{C(j,t+1,u;0,n-j+s-1;d)}{C(s,t,u;m,n;d)}\, G_{j,t+1,u}(0,n+j+s-1;d,e).$$

Now we shall prove the following

Theorem:

For any fixed a priori distribution $\lambda(q_1,q_2)$ and for any $G(m,n;d,e)$-situation with $m \geq 2$, the size of the next test group under the procedure $R^{(1)}$, does not depend on e.

Proof: If we substitute this value of $\overline{G}_{s,t,u}(m,n;d,e)$ given by the right side of (3.4.3) in the three $\overline{G}$'s occuring in (3.3.31), we get after simplification

$$(3.4.4)\,\overline{F}_{s,t,u}(m,m;d)=1+\min_{1\leq x\leq m-1}\Big\{\frac{C(s+x,t,u;m-x,n-x;d)}{C(s,t,u;m,n;d)}\,\overline{F}_{s+x,t,u}(m-x,n-x;d,e)$$

$$+\frac{C(s,t,u;x,n;d)}{C(s,t,u;m,n;d)}\,\overline{F}_{s,t,u}(x,n;d,e)\Big\},$$

which is independent of e. The boundary condition states that $\overline{F}_{s,t,u}(1,n;d) = 0$ and this also does not depend on e. It is clear from the drivation that (3.4.4), which does not depend on e, must define the same integer value x as defined by (3.3.31). This proves the theorem.

Chapter 4

The k-category Problem with Nested Dominance

4.1 Formulation of the Problem

A finite number N of units are to be classified into one of the k disjoint categories by means of group-testing. The k categories are labeled as 'the best', 'the 2^{nd}-best', ..., 'the k^{th}-best'. A group-test is a simultaneous test on x units ($1 \leq x \leq N$) with one of the following k possible outcomes: At least one of the x units belongs to the i^{th}-best category and none of the x units belongs to the j^{th}-best category for $j > i$. The problem is to define a simple and efficient procedure (or an optimal procedure) for classifying all the N units. Each unit is assumed to represent an independent observation from a multinomial population with a known a priori probability q_i of any unit belonging to the i^{th}-best category for $i=1,2,\ldots,k$ where $q_i \geq 0$ and $q_1+q_2+\ldots+q_k = 1$.

A procedure R_1 which describes a mode of action for any given value of $\vec{q} = (q_1,\ldots,q_k)$ is proposed. Under the procedure R_1, at any stage of the experiment the experimentor must separate the unclassified units into at most 2k-2 sets and units within each of these 2k-2 sets need not be distinguishable.

4.2 The Procedure R_1

Before defining the procedure R_1 we shall need some definitions and results. A set of units will be called of type $C_i (1 \leq i \leq k)$ if it is known that it contains at least one unit which belongs to the i^{th}-best category and none of the units in the set belongs to the j^{th}-best category where $j > i$. Thus a group-test tells us that a set of units is of type C_i for some $i (1 \leq i \leq k)$.

A set of units will be called of type D_j $(1 \leq j \leq k)$ if the probability that any unit in this set belongs to the i^{th}-best category, independently of any other unit in this set, is given by

$$(4.2.1) \qquad q_{i|j} = \frac{q_i}{q_{[j]}} \qquad i = 1,2,\ldots,j$$

and equals zero otherwise; here $q_{[j]} = \sum_{i=1}^{j} q_i$. At the outset all the units are of type D_k(which corresponds to the trinomial set for k = 3) and $q_{i|k} = q_i (i=1,2,\ldots,k)$.

Let Y_i $(1 \leq i \leq k)$ be the chance variable representing the number of units belonging to the i^{th}-best category in any set. Then for a set S of size $m \geq 2$ and for $i + 1 \leq j \leq k$, the conditional probability $P_{ij}(y)$ that $\sum_{a=i+1}^{j} Y_a = y$ given that S is of type C_j is

$$(4.2.2) \quad P_{ij}(y) = P\{\sum_{a=i+1}^{j} Y_a = y | Y_j \geq 1,\ Y_{j+t} = 0 \text{ for } t = 1,2,\ldots,k-j\}$$

$$= \frac{\binom{m}{y} q_{[i]}^{m-y} \sum_{t=1}^{y} \binom{y}{t} q_j^t (q_{i+1}+q_{i+2}+\ldots+q_{j-1})^{y-t}}{q_{[j]}^m - q_{[j-1]}^m}$$

$$= \frac{\binom{m}{y} q_{[i]}^{m-y} [(\sum_{a=i+1}^{j} q_a)^y - (\sum_{a=i+1}^{j-1} q_a)^y]}{q_{[j]}^m - q_{[j-1]}^m} \quad \text{for } y = 1,2,\ldots,m.$$

Then, for any fixed $i \leq j-1$, the probability that a randomly chosen proper subset of fixed size x from a set of type $C_j (2 \leq j \leq k)$ and size $m \geq 2$ is such that all its units belong to one of the first i best categories is

$$(4.2.3) \quad \sum_{y=1}^{m-x} \frac{\binom{m-y}{x}}{\binom{m}{x}} \frac{\binom{m}{y} q_{[i]}^{m-y} [(\sum_{a=i+1}^{j} q_a)^y - (\sum_{a=i+1}^{j-1} q_a)^y]}{q_{[j]}^m - q_{[j-1]}^m}$$

$$= q_{[i]}^x \sum_{y=1}^{m-x} \binom{m-x}{y} \frac{q_{[i]}^{m-y-x} [(\sum_{a=i+1}^{j} q_a)^y - (\sum_{a=i+1}^{j-1} q_a)^y]}{q_{[j]}^m - q_{[j-1]}^m}$$

$$= \frac{q_{[i]}^x [q_{[j]}^{m-x} - q_{[j-1]}^{m-x}]}{q_{[j]}^m - q_{[j-1]}^m} .$$

Similarly the probability that a randomly chosen proper subset of size x from a set of type $C_j (2 \leq j \leq k)$ of size $m \geq 2$ is such that all units belong to one of the first (i-1) best categories ($i \leq j-1$) is

$$(4.2.4) \quad \frac{q_{[i-1]}^{x}[q_{[j]}^{m-x} - q_{[j-1]}^{m-x}]}{q_{[j]}^{m} - q_{[j-1]}^{m}} .$$

Hence the probability $P_{ij}(m)$ that a randomly chosen proper subset of size x from a set of type $C_j (2 \leq j \leq k)$ of size $m \geq 2$, is of type $C_i (i \leq j-1)$ is equal to P[every element of the proper subset belongs to any one of the first i best categories] - P[every element of the proper subset belongs to any one of the first (i-1) best categories], i.e.,

$$(4.2.5) \qquad P_{ij}(m) = \frac{(q_{[i]}^{x} - q_{[i-1]}^{x})(q_{[j]}^{m-x} - q_{[j-1]}^{m-x})}{q_{[j]}^{m} - q_{[j-1]}^{m}} ,$$

where it is understood that $q_{[0]} = 0$.

It follows from (4.2.5) that for a set of type $C_j (2 \leq j \leq k)$ of size $m \geq 2$, the probability $P_{jj}(m)$ that a randomly chosen proper subset of size x is also of type C_j.

$$(4.2.6) \quad P_{jj}(m) = 1 - \sum_{i=1}^{j-1} \frac{[q_{[i]}^{x} - q_{[i-1]}^{x}][q_{[j]}^{m-x} - q_{[j-1]}^{m-x}]}{q_{[j]}^{m} - q_{[j-1]}^{m}}$$

$$= 1 - \frac{q_{[j-1]}^{x}[q_{[j]}^{m-x} - q_{[j-1]}^{m-x}]}{q_{[j]}^{m} - q_{[j-1]}^{m}}$$

$$= \frac{q_{[j]}^{m-x}[q_{[j]}^{x} - q_{[j-1]}^{x}]}{q_{[j]}^{m} - q_{[j-1]}^{m}} .$$

Now we shall prove a lemma which plays a fundamental role in the derivation of the procedure R_1.

Lemma 3: Given that a set of size $m \geq 2$ is of type C_j for any fixed j ($2 \leq j \leq k$) and given that a proper subset of size x ($1 \leq x \leq m-1$) chosen from this set also proves to be of type C_j, then the a posteriori distribution associated with the remaining set of type $(m-x)$ units is precisely the same as that of m-x units belonging to a set of type D_j, i.e., it is a multinomial distribution with index m-x and with parameters $\{q_{i|j}\}$ defined in (4.2.1).

Proof: Let A be the set of size x randomly chosen from the given set of type C_j and let B be the set of the remaining m-x units. Let A_i $(1 \leq i \leq k)$ and $B_i (1 \leq i \leq k)$ denote the random number of units of the i^{th}-best category present in A and B respectively and let $\vec{A} = (A_1,\ldots,A_k)$, $\vec{B} = (B_1,\ldots,B_k)$. Then for any j and for any set $(b_1,\ldots,b_j)$ such that $b_i \geq 0$ and $\sum_{i=1}^{j} b_i = m-x$, we have

$$(4.2.7) \quad P_j = P\{B_1=b_1,\ldots,B_j=b_j | A_j+B_j \geq 1, A_j \geq 1, A_{j+t}+B_{j+t}=0 \text{ for } t=1,\ldots,k-j\},$$

where P_j is defined by (4.2.7). Since $A_j \geq 1$ inplies $A_j+B_j \geq 1$ and $A_{j+t}+B_{j+t}=0$ implies $A_{j+t=0}$ and $B_{j+t}=0$, then

$$(4.2.8) \quad P_j = \frac{P\{B_1=b_1,\ldots,B_j=b_j, B_{j+t}=0 \text{ for } t=1,\ldots,k-j \text{ and } A_j \geq 1, A_{j+t}=0 \text{ for } t=1,\ldots,k-j\}}{P\{B_{j+t}=0 \text{ for } t=1,\ldots,k-j \text{ and } A_j \geq 1, A_{j+t}=0 \text{ for } t=1,\ldots,k-j\}} \; :$$

At the outset (and in the unconditional probability above) all the units are independently and multinomially distributed with a common probability $q_i (1 \leq i \leq k)$ of belonging to the i^{th}-best category. Since the sets A and B are disjoint it follows that $\vec{A}$ and $\vec{B}$ are independent vector chance variables. Hence both the numerator and the denominator factor and after cancellation of the second factors, we obtain

$$(4.2.9) \quad P_j = \frac{(m-x)!}{\prod_{i=1}^{j} (b_i!)} \; \frac{q_1^{b_1} q_2^{b_2} \cdots q_j^{b_j}}{(q_1+\ldots q_j)^{m-x}}$$

$$= \frac{(m-x)!}{\prod_{i=1}^{j} (b_i!)} \; \prod_{i=1}^{j} \left(\frac{q_i}{q_{[j]}}\right)^{b_i}$$

$$= \frac{(m-x)!}{\prod_{i=1}^{j} (b_i!)} \prod_{i=1}^{j} q_{i|j}^{b_i} ,$$

which proves the lemma.

Let $G_1(n_1;{}^{m_2}_{n_2};\ldots;{}^{m_k}_{n_k};\vec{q}) = G({}^{m_2}_{n_2};\ldots;{}^{m_k}_{n_k})$ denote the expected number of group-tests to be performed if presently the number of classified units is n_1, the size of the set of type $C_i(2 \leq i \leq k)$ is m_i, the size of the sets of $D_i(2 \leq i \leq k)$ is n_i-m_i, the a priori probability of a unit belonging to the i^{th}-best category is $q_i(i=1,\ldots,k)$ and the procedure R_1 is used; for the special case when $m_2=\ldots=m_k=0$ we use the symbol $H_1(n_1;n_2;\ldots;n_k;\vec{q}) = H(n_2;\ldots;n_k)$. The values of $m_i(2 \leq i \leq k)$ and $n_i(1 \leq i \leq k)$ vary as the procedure R_1 is carried out; at the outset $m_2=\ldots=m_k=n_1=\ldots=n_{k-1}=0$ and $n_k=N$. The situation when $\max(m_2,\ldots;m_k) \geq 2$ will be referred to as a G-situation and the situation when $m_2=\ldots=m_k=0$ will be referred to as an H-situation. The situation when $\max(m_2,\ldots,m_k) = 1$ is excluded in the above definition since this situation can be reduced to an H-situation without any test.

<u>Recursion Formula Defining Procedure R_1:</u>

For any H-situation (i.e., for $m_i=0;i=2,\ldots,k$) with $n_i \geq 0$ for $i=2,\ldots,k-1$ and $n_k \geq 1$, we take a sample of size $x(1 \leq x \leq n_k)$ from the size of type D_k and then

$$(4.2.10)\quad H(n_2;n_3;\ldots;n_k) = 1+\min_{1\leq x\leq n_k}\{q_1^x H(n_2;\ldots;n_{k-1};n_k-x) +$$

$$(q_{[2]}^x-q_1^x)G({}^{x}_{n_2+x};{}^{0}_{n_3};\ldots;{}^{0}_{n_k-x}) + \ldots +$$

$$(q_{[k-1]}^x-q_{[k-2]}^x)G({}^{0}_{n_2};{}^{0}_{n_3};\ldots;{}^{x}_{n_{k-1}+x};{}^{0}_{n_k-x}) +$$

$$(q_{[k]}^x-q_{[k-1]}^x)G({}^{0}_{n_2};\ldots;{}^{0}_{n_{k-1}};{}^{x}_{n})\}$$

$$= 1+\min_{1\leq x\leq n_k}\{q_1^x H(n_2;\ldots;n_{k-1};n_k-x) +$$

$$\sum_{i=2}^{k}(q^x_{[i]}-q^x_{[i-1]})G\begin{pmatrix}x\delta_{2,i} & ;\ldots; & x\delta_{k-1,i} & ; & x\delta_{k,i}\\ x\delta_{2,i}+n_2 & & x\delta_{k-1,i}+n_{k-1} & & n_k-x+x\delta_{k,i}\end{pmatrix},$$

where $\delta_{i,j}=1$ if $i=j$ and is zero if $i\neq j$. Let $q_{[i|j]}=\sum_{a=1}^{i} q_{i|j}$.

More generally for any H-situation (i.e., for $m_i=0, i=2,\ldots,k$) with $n_i \geq 0$, $i=1,2,\ldots,j-1$, $n_j \geq 1$, $n_{j+1}=\ldots=n_k=0$ we take a sample of size x $(1 \leq x \leq n_j)$ from the set of type D_j and then

$$(4.2.11)\quad H(n_2;\ldots;n_j;0,\ldots;0) = 1+\min_{1\leq x\leq n_j}\{q^x_{1|j}H(n_2;\ldots;n_j-x;0;\ldots;0)+$$

$$\sum_{i=2}^{j}\left(q^x_{[i|j]}-q^x_{[i-1|j]}\right)G\begin{pmatrix}x\delta_{2,i} & ;\ldots; & x\delta_{j-1,i} & ; & x\delta_{j,i} & ;0;\ldots;0\\ n_2+x\delta_{2,i} & & n_{j-1}+x\delta_{j-1,i} & & n_j-x+x\delta_{j,i} & \;\;0\quad\;\; 0\end{pmatrix}\}.$$

We now use formulas (4.2.5), (4.2.6) and lemma 3. For any G-situation with $m_2 \geq 2$ (and any values of $n_2 \geq m_2, n_3 \geq m_3 \geq 0,\ldots,n_k \geq m_k \geq 0$) we take a sample of size $x(1 \leq x \leq m_2-1)$ from the set of type C_2 and then

$$(4.2.12)\quad G\binom{m_2}{n_2};\ldots;\binom{m_k}{n_k} = 1+\min_{1\leq x\leq m_2-1}\{\frac{q_1^x(q_{[2]}^{m_2-x}-q_1^{m_2-x})}{q_{[2]}^{m_2}-q_1^{m_2}}G\left(\begin{matrix}m_2-x\\ n_2-x\end{matrix};\begin{matrix}m_3\\ n_3\end{matrix};\ldots;\begin{matrix}m_k\\ n_k\end{matrix}\right)$$

$$+\frac{(q^x_{[2]}-q^x_1)q_{[2]}^{m_2-x}}{q_{[2]}^{m_2}-q_1^{m_2}}\;G\left(\begin{matrix}x\\ n_2\end{matrix};\begin{matrix}m_3\\ n_3\end{matrix};\ldots;\begin{matrix}m_k\\ n_k\end{matrix}\right)\}.$$

More generally, for any G-situation with $m_2=\ldots=m_{j-1}=0, m_j \geq 2$(and any values of $n_2 \geq 0,\ldots,n_{j-1} \geq 0$, $n_j \geq m_j, n_{j+1} \geq m_{j+1} \geq 0,\ldots,n_k \geq m_k \geq 0$) we take a sample of size $x(1 \leq x \leq m_j-1)$ from the set of type C_j and then

$$(4.2.13)\quad G\left(\begin{matrix}0\\ n_2\end{matrix};\ldots;\begin{matrix}0\\ n_{j-1}\end{matrix};\begin{matrix}m_j\\ n_j\end{matrix};\ldots;\begin{matrix}m_k\\ n_k\end{matrix}\right) = 1+$$

$$\min_{1\leq x\leq m_2-1}\{\sum_{i=1}^{j-1}P_{ij}(m_j)G\left(\begin{matrix}x\delta_{2,i}\\ n_2+x\delta_{2,i}\end{matrix};\ldots;\begin{matrix}x\delta_{j-1,i}\\ n_{j-1}+x\delta_{j-1,i}\end{matrix};\begin{matrix}m_j-x\\ n_j-x\end{matrix};\begin{matrix}m_{j+1}\\ n_{j+1}\end{matrix};\ldots;\begin{matrix}m_k\\ n_k\end{matrix}\right)$$

$$+ P_{jj}(m_j)G\left({0 \atop n_2};\ldots;{0 \atop n_{j-1}};{x \atop n_j};{m_{j+1} \atop n_{j+1}};\ldots;{m_k \atop n_k}\right)\},$$

where $P_{ij}(\cdot)$ and $P_{jj}(\cdot)$ are defined in (4.2.5) and (4.2.6).

The boundary conditions state that for all $\vec{q} = (q_1,\ldots,q_k)$

$$H(0;0;\ldots;0) = 0.$$

$$G\left({1 \atop n_2};{m_3 \atop n_3};\ldots;{m_k \atop n_k}\right) = G\left({0 \atop n_2-1};{m_3 \atop n_3};\ldots;{m_k \atop n_k}\right) \text{ for } n_2 \geq 1,\ n_3 \geq m_3 \geq 0,\ldots,n_k \geq m_k \geq 0.$$

$$G\left({0 \atop n_2};{1 \atop n_3};{m_4 \atop n_4};\ldots;{m_k \atop n_k}\right) = G\left({0 \atop n_2};{0 \atop n_3-1};{m_4 \atop n_4};\ldots;{m_k \atop n_k}\right) \text{ for } n_2 \geq 0, n_3 \geq 1, n_4 \geq m_4 \geq 0,\ldots,n_k \geq m_k \geq 0.$$

(4.2.14)

etc.

$$G\left({0 \atop n_2};\ldots;{0 \atop n_{k-1}};{1 \atop n_k}\right) = H(n_2;n_3;\ldots;n_{k-1};n_k-1) \text{ for } n_2 \geq 0,\ldots,n_{k-1} \geq 0, n_k \geq 1.$$

and for any $j (j=1,2,\ldots,k)$

$$G\left({0 \atop n_2};\ldots;{0 \atop n_{j-1}};{1 \atop n_j};{0 \atop n_{j+1}};\ldots;{0 \atop n_k}\right) = H(n_2;\ldots;n_{j-1};n_j-1;n_{j+1};\ldots;n_k) \text{ for}$$

$n_2;n_3,\ldots,n_{j-1},n_{j+1},\ldots,n_k \geq 0, n_j \geq 1.$

4.3 Properties of the Procedure R_1

In this section we consider a property of the procedure R_1 which is concerned with the size of the next test group when a set of type C_2 is of size $m_2 \geq 2$. This property is a generalization of the corresponding result in section 1.4 for the 3-category problem. We state this property as a

Theorem:

For any $G\left({m_2 \atop n_2};{m_3 \atop n_3};\ldots;{m_k \atop n_k}\right)$ -situation with $m_2 \geq 2$ and any $\vec{q} = (q_1,q_2,\ldots,q_k)$, the size of the next group-test, under procedure R_1, depends only on m_2 (and does not depend on n_2 or m_3 or $n_3,\ldots,$ or m_k or n_k)

Proof: For this situation the procedure under R_1 is to break down the set of type C_2 until a single unit belonging to 2^{nd}-best category is found and removed.

(Instead of randomizing the units in this set each time before a test group is selected, it is assumed, without any loss of generality, the order is randomized only once at the outset; units removed for testing will be taken in that order.) If the ith unit is the first unit belonging to 2^{nd}-best category, then the breaking down of the set of type C_2 leads to a situation in which the set of type C_2 is empty and the size of the set of type D_2 is increased by m_2-i whereas the sizes of the other sets remain the same; i.e., it leads to $G\binom{0}{n_2-i};\binom{m_3}{n_3};\ldots;\binom{m_k}{n_k}$ situation.

Let $F(m_2;\vec{q}) = F(m_2)$ be defined as the expected number of group tests required to break down the set of type $C_2$of size m_2 and reach (for the first time) a situation in which the set of type C_2 is empty, when $\vec{q}$ is given and the procedure R_1 is used. It follows from this definition that $F(m_2)$ does not depend on n_2 or m_3 or n_3 or ... m_k or n_k. Using q for $q_{1|2}$, we have

$$(4.3.1)\quad G\left(\begin{smallmatrix}m_2\\n_2\end{smallmatrix};\ldots;\begin{smallmatrix}m_k\\n_k\end{smallmatrix}\right) = F(m_2)+ \sum_{i=1}^{m} \frac{(1-q)q^{i-1}}{1-q^m}\; G\left(\begin{smallmatrix}0\\n_2-i\end{smallmatrix};\begin{smallmatrix}m_3\\n_3\end{smallmatrix};\ldots;\begin{smallmatrix}m_k\\n_k\end{smallmatrix}\right) .$$

Let $\left(\frac{1-q^m}{1-q}\right)F(m_2) = F^*(m_2)$ and $\left(\frac{1-q^{m_2}}{1-q}\right) G\left(\begin{smallmatrix}m_2\\n_2\end{smallmatrix};\ldots;\begin{smallmatrix}m_k\\n_k\end{smallmatrix}\right) = G^*\left(\begin{smallmatrix}m_2\\n_2\end{smallmatrix};\ldots;\begin{smallmatrix}m_k\\n_k\end{smallmatrix}\right)$.

Replacing F by F* and G by G* in (4.2.12) and (4.3.1), we obtain

$$(4.3.2)\quad G^*\left(\begin{smallmatrix}m_2\\n_2\end{smallmatrix};\ldots;\begin{smallmatrix}m_k\\n_k\end{smallmatrix}\right) = F^*(m_2)+ \sum_{i=1}^{m_2} q^{i-1}\; G^*\left(\begin{smallmatrix}0\\n-i\end{smallmatrix};\begin{smallmatrix}m_3\\n_3\end{smallmatrix};\ldots;\begin{smallmatrix}m_k\\n_k\end{smallmatrix}\right)$$

$$(4.3.3)\quad G^*\left(\begin{smallmatrix}m_2\\n_2\end{smallmatrix};\ldots;\begin{smallmatrix}m_k\\n_k\end{smallmatrix}\right) = \sum_{i=1}^{m_2} q^{i} + \min_{1\le x<m_2}\left\{q^x G^*\left(\begin{smallmatrix}m_2-x\\n_2-x\end{smallmatrix};\begin{smallmatrix}m_3\\n_3\end{smallmatrix};\ldots;\begin{smallmatrix}m_k\\n_k\end{smallmatrix}\right)+G^*\left(\begin{smallmatrix}x\\n_2\end{smallmatrix};\begin{smallmatrix}m_3\\n_3\end{smallmatrix};\ldots;\begin{smallmatrix}m_k\\n_k\end{smallmatrix}\right)\right\}.$$

Substituting (4.3.2) in (4.3.3) and observing that summation terms cancel, we obtain

$$(4.3.4)\quad F^*(m_2) = \frac{1-q^{m_2}}{1-q} + \min_{1\le x<m_2}\{q^x F^*(m_2-x)+F^*(x)\}$$

which does not depend on n_2 or m_3 or ... or m_k or n_k. The boundary condition is $F^*(1) = 0$ for all $\vec{q}$, which also does not depend on n_2 or m_3 or n_3 or ... or n_k. It is clear from the derivation that (4.3.4), which does not depend on n_2 or m_3 or ... or n_k, must define the same integer value x as defined by (4.2.12).

It follows from (4.3.4) that for any G-situation with $m_2 \geqq 2$ we can use the G-tables given for the binomial problem in [5].

Under the procedure R_1 for any k it is interesting to note that once a unit is classified it is never used in subsequent tests.

4.4 Some Properties of the Optimal Procedure R_0

In this section we discuss some properties of the optimal procedure R_0 for the k-category problem with nested dominance. These properties are concerned with the question of when we should test one at a time.

A group-testing procedure is called optimal among all procedures for given N and $\vec{q} = (q_1,\ldots,q_k)$ if it minimizes the expected number of tests. Let R_0 denote the optimal group-testing procedure and let $E(T|R_0,N,\vec{q}) = ET$ denote the expected number of group-tests to be performed if the procedure R_0 is used for given $\vec{q} = (q_1,\ldots,q_k)$ and starting with N units. The notation $q_{[j]}$ is the same as in section 4.1.

The optimal procedure R_0 for N = 1 is, of course, trivial and we now consider R_0 for $N \geqq 2$.

Theorem:

For $k \geqq 4$ and $N \geqq 2$ the optimal procedure R_0 has the following properties:

(i) If

$$\sum_{j=k-1}^{k} q_j q_{[j]} + \sum_{j=2}^{k-2} q_j(q_1+q_j) > q_1^2 \tag{4.4.1}$$

then ET = N and the units are all tested one at a time.

(ii) If

$$\sum_{j=2}^{k} q_j q_{[j]} < q_1^2 \tag{4.4.2}$$

then ET < N and at least one test is made with more than one unit.

Proof: The proof of statement (i) is similar but not exactly the same as the proof of the corresponding result of the theorem in section 2.1. To prove (i) we go through the details of setting up a "new" plan and comparing it with "old" plan; the notation G and B are the same as in section 2.1. Suppose that G has 3 or more units and that $k \geq 4$. Let the element u of G be any unit of G except that, by assumption, if G is of C_{k-1}-type or C_k-type, then u is not the last one among the elements of G to be tested under the old plan. If G has only 2 units then u be any unit in G. Instead of testing G at the point B, under the new plan we test G-u. If G-u is found of type C_k we continue as under the old plan where G is found to be of type C_k. If G-u is of type C_i $(1 \leq i \leq k-1)$, we test the unit u on the next test. After testing u we continue testing as we would have done under the old plan after observing the nature of the set G except that we will skip tests on units of G whose result can be inferred due to availability of the additional information. We discuss the two cases $x = 2$ and $x \geq 3$ separately:

Case I: $x = 2$

When G is of type C_i and the first unit to be tested after G under the old plan belongs to the i^{th}-best category $(2 \leq i \leq k)$ we have a saving of one test for the new plan.

In addition we would be using two tests under the new plan when G is best (i.e. of type C_1) whereas we would have used only one test under the old plan (i.e. we have a loss of one test). Thus the expected number of tests saved under the new plan is

$$\sum_{j=2}^{k} q_j q_{[j]} - q_1^2 .$$

Hence we have a positive saving when

$$\sum_{j=2}^{k} q_j q_{[j]} - q_1^2 > 0. \tag{4.4.3}$$

Case II: $x \geq 3$

Under the new plan we will have a saving of x-2 (or x-1) tests when all the x-1 units in G except u are best and u is of type C_i for $2 \leq i \leq k$. Also we shall save one test in the following situations:

(i) u is "k^{th}-best" and the unit of G-u to be tested last is "i^{th}-best"

$(2 \leq i \leq k)$ whereas all other units of G are best. (ii) u is "$(k-1)^{st}$-best" and the unit of $G-u$ to be tested last is "i^{th}-best" $(2 \leq i \leq k-1)$ whereas all other units of G are best. (iii) u and the unit of $G-u$ to be tested last are "i^{th}-best" for $2 \leq i \leq k-2$, whereas all other units of G are best.

Finally if all units of G are best we use two tests under the new plan whereas we would have used only one test under the old plan; hence there is a loss of exactly one test under the new plan.

Combining all the above results we find that the expected number of tests saved (which we denoted by S) satisfies the follwoing inequality

$$(4.4.4) \qquad S \geq q_1^x \left(\frac{q_k \sum_{i=2}^{k} q_i + q_{k-1} \sum_{i=2}^{k-1} q_i + \sum_{i=2}^{k-2} q_i^2}{q_1^2} + (x-2) \frac{\sum_{i=2}^{k} q_i}{q_1} - 1 \right) .$$

We find that the right hand side of (4.4.4) is positive for all $x \geq 3$ if it is positive for $x = 3$, i.e., if

$$(4.4.5) \qquad q_k \sum_{i=2}^{k} q_i + q_{k-1} \sum_{i=2}^{k-1} q_i + \sum_{i=2}^{k-2} q_i^2 + q_1 \sum_{i=2}^{k} q_i - q_1^2 > 0 .$$

The above result also holds for $k = 3$ if the third term is removed.

Since the difference between the left side of (4.4.3) and the left side of (4.4.5) is positive, therefore the set of $\vec{q} = (q_1, \ldots, q_k)$ which satisfies (4.4.5) also satisfies (4.4.3). Thus there is a positive saving in the expected number of tests for any $x \geq 2$ under the new plan when (4.4.5) holds, or equivalently when (4.4.1) holds, and the point B is reached. Furthermore it is evident as explained in section 2.1 that the samples with the above mentioned cases will reach the point B. Thus the number of expected tests under the new procedure is less than the expected number of tests under the old procedure whenever (4.4.1) holds. This proves the statement (i) of our theorem.

The proof of the statement (ii) of this theorem is similar to that of statement (ii) of the theorem in section 2.1 and we will only sketch the proof here. If $N = 2$, then by direct computation

$$ET \begin{cases} = 2 + \sum_{j=2}^{k} q_j q_{[j]} - q_1^2 < 2 & \text{whenever (4.4.2) holds} \\ = 2 & \text{whenever (4.4.2) does not hold.} \end{cases}$$

In the first case the optimal procedure starts by testing 2 units, in the second case it tests each unit separately.

If $N = 2M$ (any even number) we divide N units into M subgroups each of size 2 and use the optimal procedure for each subgroup of size 2. Then the expected number of tests E_0T under this scheme is

$$E_0T = M[2 + \sum_{j=2}^{k} q_j q_{[j]} - q_1^2] < 2M = N \quad \text{whenever (4.2.2) holds.}$$

Likewise for $N = 2M + 1$ we divide these units into $(M+1)$ subgroups, of which M are of size 2 and 1 is of size 1. Following the optimal procedure for each subgroup, we observe that the expected number of tests under this scheme is less than N whenever (4.4.2) holds. This proves the statement (ii) and completes the proof of our theorem.

It may be pointed out that for $k = 4$ (as well as for $k = 2$ and 3) the left sides of the inequalities in (4.4.1) and (4.4.2) are the same. Hence in these cases we know exactly for which $\vec{q}$ the optimal procedure tells us to test one at a time and for which it requires at least one group test of size at least 2. For $k \geq 5$, we do not know exactly the curve which determines the region for one at a time testing, although we know two curves between which it lies; these are given by putting an equality sign in (4.4.1) and (4.4.2).

Corollary: For $q_1 < \frac{1}{2}$, the optimal procedure (independently of k) tests one unit at a time.

Proof: We know from the result of [8] and the corollary to the theorem in section 2.1, that for $k = 2$ and 3, this corollary is true. Therefore we need consider only $k \geq 4$. The inequality (4.4.1) is always satisfied when

$$\sum_{j=2}^{k} q_j^2 + q_1 \left(\sum_{j=2}^{k} q_j \right) = q_1^2$$

or

$$(4.4.6) \qquad \sum_{j=2}^{k} q_j^2 + q_1(1-q_1) - q_1^2 = 0.$$

This curve (if q_1 is the vertical coordinate) will have a minimum value when $q_2 = \ldots = q_k = \frac{1-q_1}{k-1}$. Substituting $q_2 = \ldots = q_k = \frac{1-q_1}{k-1}$ in (4.4.6) we obtain

$$\frac{(1-q_1)^2}{k-1} + q_1 - 2q_1^2 = 0$$

or

$$(4.4.7) \qquad q_1^{(0)} = \frac{\sqrt{k^2+2k-3} + k-3}{2(2k-3)} .$$

To show that $q_1^{(0)}$ is a decreasing function of k, we find that

$$\frac{dq_1^{(0)}}{dk} = \frac{3\sqrt{k^2+2k-3} - 5k+3}{2(2k-3)^2\sqrt{k^2+2k-3}}$$

which is less than zero for $k \geq 4$. Now the right side of (4.4.7) $\to \frac{1}{2}$ from above as $k \to \infty$. The curve given by (4.4.6) is everywhere under (if q_1 is vertical coordinate) the curve corresponding to (4.4.1) with equality. Since this new curve approaches a minimum value $q_1^{(0)} \to \frac{1}{2}$ from above, it follows that the curve corresponding to equality in (4.4.1) is everywhere above the plane $q_1 = \frac{1}{2}$ in k-1 dimensional space. This proves the corollary.

We know that the optimal procedure is to test one at a time for $q_1 < \frac{\sqrt{5}-1}{2} = .618 \ldots$ for k = 2 which is consistent with (4.4.7); for k = 3 it is shown in section 2.1 that the optimal procedure tests one at a time for $q_1 < .6$ and (4.4.7) gives $\sqrt{3}/3 = .577 \ldots$. It is reasonable to conjecture (but we have not proved) that these cut off points which separate multiple group-testing from testing individual units which we know to be .618... for k = 2, .6 for k = 3) form a decreasing sequence $\to \frac{1}{2}$ as $k \to \infty$.

Chapter 5

Modified Problems related to the 3-Category Problem and Procedures for them

5.1 Problem involving two types of units

Suppose we have two easily distinguishable types of units which can be put in the same test group for the purpose of classifying each of these units into one of the three disjoint categories. Each unit of type 1 is assumed to represent an independent observation from the trinomial population with a common known probability, p_1 of being good, p_2 of being mediocre and $p_3 = 1-p_1-p_2$ of being defective; similarly for type 2 units the corresponding values are q_1, q_2 and $q_3 = 1-q_1-q_2$. The problem is to devise a procedure which classifies all the units into one of the three disjoint categories.

Let $H_{11}(n_1,n_2;e_1,e_2)$ denote the expected number of group tests required under the procedure R_{11} (to be described below) if currently the conditional binomial set contains n_i units of type i ($i = 1,2$), the trinomial set contains e_i units of type i ($i = 1,2$) and the defective and mediocre sets are empty. Let $G_{11}(m_1,m_2;n_1,n_2;d_1,d_2;e_1,e_2)$ denote the expected number of group tests required under R_{11} if the mediocre set contains m_1 units of type 1 and m_2 units of type 2 (the combined set of m_1+m_2 units is known to contain at least one mediocre unit and no defective units), the conditional binomial set contains n_1-m_1 units of type 1 and n_2-m_2 units of type 2, the defective set contains d_1 units of type 1 and d_2 units of type 2 (the combined set of d_1+d_2 units is known to contain at least one defective unit) and a trinomial set containing e_1-d_1 units of type 1 and e_2-d_2 units of type 2. The notation H-situation and G-situation is the same as in chapter 1.

Let $p_{[2]} = p_1+p_2, \quad p = p_1/p_{[2]}$,

$q_{[2]} = q_1+q_2, \quad q = q_1/q_{[2]}$.

For any H-situation with $e_1+e_2 \geq 1$, $n_1+n_2 \geq 0$ (and $m_1 = m_2 = d_1 = d_2 = 0$) we take a sample of x units of type 1 and y units of type 2 from the trinomial set and we then have

$$(5.1.1)\qquad H_{11}(n_1,n_2;e_1,e_2) = 1+\min\{p_1^x q_1^y H_{11}(n_1,n_2;e_1-x,e_2-y)$$
$$+(p_{[2]}^x q_{[2]}^y - p_1^x q_1^y)G_{11}(x,y;n_1+x,n_2+y;0,0;e_1-x,e_2-y)$$
$$+(1-p_{[2]}^x q_{[2]}^y)G_{11}(0,0;n_1,n_2;x,y;e_1,e_2)\}$$

where the minimization is carried over pairs (x,y) with $0 \leq x \leq e_1$, $0 \leq y \leq e_2$ and $x+y \geq 1$.

For any H-situation with $e_1 = e_2 = 0$, $n_1+n_2 \geq 1$, we take a sample of x units of type 1 and y units of type 2 from the conditional binomial set and we then have

$$(5.1.2)\qquad H_{11}(n_1,n_2;0,0) = 1+\min\{p^x q^y H_{11}(n_1-x,n_2-y;0,0) +$$
$$(1-p^x q^y)G(x,y;n_1,n_2;0,0;0,0)\}$$

where the minimization is carried over the pairs (x,y) with $0 \leq x \leq n_1$, $0 \leq y \leq n_2$ and $x+y \geq 1$.

For any G-situation with $m_1+m_2 \geq 2$, we take a sample of x units of type 1 and y units of type 2 from the mediocre set and letting $\vec{m} = (m_1,m_2)$, $\vec{n} = (n_1,n_2)$, $\vec{d} = (d_1,d_2)$ and $\vec{e} = (e_1,e_2)$ we then have

$$(5.1.3)\qquad G_{11}(\vec{m};\vec{n};\vec{d};\vec{e}) = 1+\min\left\{ \frac{p^x q^y - p^{m_1} q^{m_2}}{1-p^{m_1} q^{m_2}} G_{11}(m_1-x,m_2-y;n_1-x,n_2-y;\vec{d};\vec{e}) + \right.$$
$$\left. \frac{1-p^x q^y}{1-p^{m_1} q^{m_2}} G_{11}(x,y;\vec{n};\vec{d};\vec{e}) \right\}$$

where the minimization is carried over the pairs (x,y) with $0 \leq x \leq m_1$, $0 \leq y \leq m_2$ and $1 \leq x+y \leq m_1+m_2-1$.

For any G-situation with $m_1 = m_2 = 0$ and $d_1+d_2 \geq 2$, we take a sample of x units of type 1 and y units of type 2 from the defective set and we then have

$$(5.1.4)\qquad G_{11}(0,0;\vec{n};\vec{d};\vec{e}) = 1+\min\left\{ \frac{p_1^x q_1^y(1-p_{[2]}^{d_1-x} q_{[2]}^{d_2-y})}{1-p_{[2]}^{d_1} q_{[2]}^{d_2}}\, G_{11}(0,0;\vec{n};d_1-x,d_2-y;e_1-x,e_2-y)\right.$$

$$+ \frac{(p_{[2]}^x q_{[2]}^y - p_1^x q_1^y)(1-p_{[2]}^{d_1-x} q_{[2]}^{d_2-y})}{1-p_{[2]}^{d_1} q_{[2]}^{d_2}}\, G_{11}(x,y;n_1+x,n_2+y;d_1-x,d_2-y;e_1-x,e_2-y)$$

$$\left. + \frac{1-p_{[2]}^x q_{[2]}^y}{1-p_{[2]}^{d_1} q_{[2]}^{d_2}}\, G_{11}(0,0;\vec{n};x,y;\vec{e})\right\}$$

where the minimization is carried over the pairs (x,y) with $0 \leqq x \leqq d_1$, $0 \leqq y \leqq d_2$ and $0 \leqq x+y \leqq d_1+d_2-1$. The case of G-situation with $m_1+m_2 = 1$ and $d_1+d_2 = 1$ is taken care of by the boundary conditions given below.

The boundary conditions state that for all vectors $\vec{p} = (p_1,p_2,p_3)$ and $\vec{q} = (q_1,q_2,q_3)$

(5.1.5) $H_{11}(n_1,0;e_1,0) = H_1(n_1;e_1;\vec{p})$.

(5.1.6) $H_{11}(0,n_2;0,e_2) = H_1(n_2;e_2;\vec{q})$.

(5.1.7) $G_{11}(m_1,0;n_1,0;d_1,0;e_1,0) = G_1(m_1,n_1;d_1,e_1;\vec{p})$.

(5.1.8) $G_{11}(0,m_2;0,n_2;0,d_2;0,e_2) = G_1(m_2,n_2;d_3,e_2;\vec{q})$.

(5.1.9) $G_{11}(1,0;n_1,n_2;0,0;\vec{e}) = H_{11}(n_1-1,n_2;\vec{e})$ for all $n_1 \geqq 1$, $n_2 \geqq 0$, $e_1 \geqq 0$, $e_2 \geqq 0$.

(5.1.10) $G_{11}(0,1;n_1,n_2;0,0;\vec{e}) = H_{11}(n_1 n_2-1;\vec{e})$ for all $n_1 \geqq 0$, $n_2 \geqq 1$, $e_1 \geqq 0$, $e_2 \geqq 0$.

(5.1.11) $G_{11}(0,0;n_1,n_2;1,0;\vec{e}) = H_{11}(n_1,n_2;e_1-1,e_2)$ for all $n_1 \geqq 0$, $n_2 \geqq 0$, $e_1 \geqq 1$, $e_2 \geqq 0$.

(5.1.12) $G_{11}(0,0;n_1,n_2;0,1;\vec{e}) = H_{11}(n_1,n_2;e_1,e_2-1)$ for all $n_1 \geqq 0$, $n_2 \geqq 0$, $e_1 \geqq 0$, $e_2 \geqq 1$.

where the right hand members of (5.1.5) to (5.1.8) refer to the basic procedure R_1 defined earlier.

5.2 Problem involving two experimenters

Suppose there are two experimenters who are working on a single set of N units by carrying out simultaneously group tests and cooperating in such a way as to minimize the time required to classify all the N units. There will be no saving in the expected number of group tests by having two experimenters. However if we regard the simultaneous tests as a stage, each of the simultaneous tests lasting the same amount of time, then the problem of minimizing the expected time for the classification of these N units is equivalent to the problem of reducing the expected number of stages. Thus in this problem the main emphasis is to reduce the expected time for classifying all the N units.

At any stage of the experimentation there are at most two mediocre sets and at most two defective sets. Let m_1, m_2 denote the sizes of the mediocre sets and let $n-(m_1+m_2) \geq 0$ denote the size of the conditional binomial set; let d_1, d_2 denote the sizes of the defective sets and let $e-(d_1+d_2) \geq 0$ denote the size of the trinomial set; it will not matter which one is called m_1 and which one is called m_2 and the same is true for the two defective sets of sizes d_1 and d_2. Let $G(m_1,m_2,n;d_1,d_2,e)$ denote the expected number of stages required when we have two mediocre sets of sizes m_1, m_2, a conditional binomial set of size $n-(m_1+m_2)$, two defective sets of sizes d_1, d_2 and a trinomial set of size $e-(d_1+d_2)$ and the procedure R_{12} (based on the recursion formula given below) is used. For the case $m_1 = m_2 = d_1 = d_2 = 0$ we shall write $G(0,0,n;0,0,e) = H(n;e)$. Also we shall write $G(m,0,n;d_1,d_2,e) = G(m,n;d_1,d_2,e)$ and likewise expressions for similar situations. The notation H-situation and G-situation is similar to the corresponding situations in chapter 1. Also the notation $\vec{q} = (q_1,q_2,q_3)$ and $q = q_1/q_{[2]}$ is the same as in chapter 1.

Procedure R_{12}

For any H-situation with $e \geq 2$, we take two samples of sizes x and y (one for each experimenter) from the trinomial set; hence

$$(5.2.1)\quad H(n;e) = 1+ \min_{\substack{x,y\geq 1\\ x+y\leq e}} [q_1^x\{q_1^y H(n;e-x-y)+(q_{[2]}^y-q_1^y)G(y,n+y;0,e-x-y)$$
$$+ (1-q_{[2]}^y)G(0,n;y,e-x)\} +$$
$$(q_{[2]}^x-q_1^x)\{q_1^y G(x,n+x;0,e-x-y)+(q_{[2]}^y-q_1^y)G(x,y,n+x+y;0,e-x-y)$$
$$+ (1-q_{[2]}^y)G(x,n+x;y,e-x)\} +$$
$$(1-q_{[2]}^x)\{q_1^y G(0,n;x,e-y)+(q_{[2]}^y-q_1^y)G(y,n+y;x,e-y)$$
$$+ (1-q_{[2]}^y)G(0,n;x,y,e)\}].$$

For any H-situation with $e = 1$ and $n \geq 1$, we take a sample of size one from the trinomial set and a sample of size x from the conditional binomial set; hence

$$(5.2.2)\qquad H(n;1) = 1+ \min_{1\leq x\leq n} [q^x H(n-x;0) + (1-q^x)G(x,n;0,0)].$$

For the H-situation with $e = 1$ and $n = 0$ the procedure is clear and $H(0;1) = 1$.

For any H-situation with $e = 0$ and $n \geq 2$ we take two samples of sizes x and y (one sample for each experimenter) from the conditional binomial set; hence

$$(5.2.3)\quad H(n;0) = 1+ \min_{\substack{x,y\geq 1\\ x+y\leq n}} [q^x\{q^y H(n-x-y;0) + (1-q^y)G(y,n-x;0,0)\} +$$
$$(1-q^x)\}q^y H(x,n-y;0,0) + (1-q^y)G(x,y,n;0,0)\}].$$

For the H-situation with $e = 0$ and $n = 1$ the procedure is clear and $H(1;0) = 1$.

For any H-situation with $\min(m_1,m_2) \geq 2$, we take a sample of size x from one mediocre set and a sample of size y from the other mediocre set; hence

$$(5.2.4)\quad G(m_1,m_2,n;d_1,d_2,e) = 1+ \min_{\substack{1\leq x< m_1\\ 1\leq y< m_2}} \left(\frac{q^x-q^{m_1}}{1-q^{m_1}}\right)\left\{\left(\frac{q^y-q^{m_2}}{1-\,^{m_2}}\right) G(m_1-x,m_2-y,n-x-y;d_1,d_2,e)\right.$$
$$\left. + \left(\frac{1-q^y}{1-q^{m_2}}\right) G(m_1-x,y,n-x;d_1,d_2,e)\right\} +$$
$$\left(\frac{1-q^x}{1-q^{m_1}}\right)\left\{\left(\frac{q^y-q^{m_2}}{1-q^{m_2}}\right) G(x,m_2-y,n-y;d_1,d_2,e)\right.$$

$$\left(\frac{(q^{y}_{[2]}-q^{y}_{1})(1-q^{d_2-y}_{[2]})}{1-q^{d_2}_{[2]}}\right) G(y,n+y;x,d_2-y,e-y) \quad +$$

$$\left.\left.\left(\frac{1-q^{y}_{[2]}}{1-q^{d_2}_{[2]}}\right) G(0,0,n;x,y,e)\right\}\right] .$$

For any G-situation with $m_1 = 0$, $m_2 = m \geq 2$, $d_1 = 0$, $d_2 = d \geq 2$, we take two samples of sizes x and y, either both from the non-empty mediocre set or one (of size x) from the non-empty mediocre set and other (of size y) from the non-empty defective set; hence

$$(5.2.6) \qquad G(m,n;d,e) \; = \; 1 + \min \left\{ \min_{\substack{x,y\geq 1\\ x+y\leq m}} G^{(1)}(x,y), \quad \min_{\substack{1\leq x\leq m\\ 1\leq y\leq d}} G^{(2)}(x,y) \right\}$$

where

$$(5.2.7) \qquad G^{(1)}(x,y) \; = \; \left(\frac{q^{x+y}-q^{m}}{1-q^{m}}\right) G(m-x-y,n-x-y;d,e) \quad +$$

$$\left(\frac{q^{x}(1-q^{y})}{1-q^{m}}\right) G(y,n-x;d,e) \; + \left(\frac{q^{y}(1-q^{x})}{1-q^{m}}\right) G(x,n-y;d,e)$$

$$+ \left(\frac{(1-q^{x})(1-q^{y})}{1-q^{m}}\right) G(x,y,n;d,e)$$

and

$$(5.2.8) \qquad G^{(2)}(x,y) \; = \; \left(\frac{q^{x}-q^{m}}{1-q^{m}}\right)\left\{\left(\frac{q^{y}_{1}(1-q^{d-y}_{[2]})}{1-q^{d}_{[2]}}\right) G(m-x,n-x;d-y,e-y)\right.$$

$$+ \; \frac{(q^{y}_{[2]}-q^{y}_{1})(1-q^{d-y}_{[2]})}{1-q^{d}_{[2]}} \quad G(m-x,y,n-x+y;d-y,e-y)$$

$$+ \left(\frac{1-q^{y}}{1-q^{m_2}}\right) G(x,y,n;d_1,d_2,e)\Bigg\} .$$

For any G-situation with $m_1 = m_2 = 0$ and $\min(d_1,d_2) \geq 2$ we take a sample of size x from one defective set and a sample of size y from the other defective set; hence

$$(5.2.5)\quad G(0,0,n;d_1,d_2,e) = 1+ \min_{\substack{1\leq x<d_1\\ 1\leq y<d_2}} \Bigg[\left(\frac{q_1^{x}(1-q_{[2]}^{d_1-x})}{1-q_{[2]}^{d_1}}\right)\Bigg\{\left(\frac{q_1^{y}(1-q_{[2]}^{d_2-y})}{1-q_{[2]}^{d_2}}\right) G(0,0,n;d_1-x,d_2-y,e-x-y)$$

$$+ \left(\frac{(q_{[2]}^{y}-q_1^{y})(1-q_{[2]}^{d_2-y})}{1-q_{[2]}^{d_2}}\right) G(y,n+y;d_1-x,d_2-y,e-x-y)$$

$$+ \left(\frac{1-q_{[2]}^{y}}{1-q_{[2]}^{d_2}}\right) G(0,0,n;d_1-x,y,e-x)\Bigg\} +$$

$$\left(\frac{(q_{[2]}^{x}-q_1^{x})(1-q_{[2]}^{d_1-x})}{1-q_{[2]}^{d_1}}\right)\Bigg\{\left(\frac{q_1^{y}(1-q_{[2]}^{d_2-y})}{1-q_{[2]}^{d_2}}\right) G(x,n+x;d_1-x,d_2-y,e-x-y)$$

$$+ \left(\frac{(q_{[2]}^{y}-q_1^{y})(1-q_{[2]}^{d_2-y})}{1-q_{[2]}^{d_2}}\right) G(x,y,n+x+y;d_1-x,d_2-y,e-x-y)$$

$$+ \left(\frac{1-q_{[2]}^{y}}{1-q_{[2]}^{d_2}}\right) G(x,n+x;d_1-x,y,e-x)\Bigg\} +$$

$$\left(\frac{1-q_{[2]}^{x}}{1-q_{[2]}^{d_1}}\right)\Bigg\{\left(\frac{q_1^{y}(1-q_{[2]}^{d_2-y})}{1-q_{[2]}^{d_2}}\right) G(0,0,n;x,d_2-y,e-y) +$$

$$+\left(\frac{1-q_{[2]}^{y}}{1-q_{[2]}^{d}}\right) G(m-x,n-x;y,e)\Bigg\} +$$

$$\left(\frac{1-q^{x}}{1-q^{m}}\right)\left\{\left(\frac{q_1^{y}(1-q_{[2]}^{d-y})}{1-q_{[2]}^{d}}\right) G(x,n;d-y,e-y) \;+\right.$$

$$\left.\left(\frac{(q_{[2]}^{y}-q_1^{y})(1-q_{[2]}^{d-y})}{1-q_{[2]}^{d}}\right) G(x,y,n+y;d-y,e-y) + \left(\frac{1-q_{[2]}^{y}}{1-q_{[2]}^{d}}\right) G(x,n;y,e)\right\} .$$

For any G-situation with $m_1 = m_2 = 0$, $d_1 = 0$, $d_2 = d \geq 2$, $e > d$, we take two samples of sizes x and y, either both from the non-empty defective set or one sample of size x from the non-empty defective set and one sample of size y from the trinomial set; hence

$$(5.2.9)\qquad G(0,n;d,e) = 1 + \min\left\{\min_{\substack{x,y\geq 1\\ x+y<d}} G^{(3)}(x,y),\quad \min_{\substack{1\leq x<d\\ 1\leq y\leq e-d}} G^{(4)}(x,y)\right\}$$

where

$$(5.2.10)\qquad G^{(3)}(x,y) = \left(\frac{q_1^{x}}{1-q_{[2]}^{d}}\right)\{q_1^{y}(1-q_{[2]}^{d-x-y})G(0,n;d-x-y,e-x-y)$$

$$+ (q_{[2]}^{y}-q_1^{y})(1-q_{[2]}^{d-x-y})G(y,n+y;d-x-y,e-x-y)$$

$$+ (1-q_{[2]}^{y})G(0,n;y,d-x)\} \;+$$

$$\left(\frac{q_{[2]}^{x}-q_1^{x}}{1-q_{[2]}^{d}}\right)\{q_1^{y}(1-q_{[2]}^{d-x-y})G(x,n+x;d-x-y,e-x-y) \;+$$

$$(q_{[2]}^{y}-q_1^{y})(1-q_{[2]}^{d-x-y})G(x,y,n+x+y;d-x-y,e-x-y)$$

$$+ (1-q_{[2]}^{y})G(x,n+x;y,e-x)\} \;+$$

$$\left(\frac{1-q_{[2]}^{x}}{1-q_{[2]}^{d}}\right)\{q_1^y G(0,n;x,e-y) + (q_{[2]}^{y}-q_1^y)G(y,n+y;x,e-y)$$

$$+ (1-q_{[2]}^{y})G(0,n;x,y,e)\}$$

and

$$(5.2.11) \quad G^{(4)}(x,y) = \left(\frac{q_1^x(1-q_{[2]}^{d-x})}{1-q_{[2]}^{d}}\right)\{q_1^y G(0,n;d-x,e-x-y) +$$

$$(q_{[2]}^{y}-q_1^y)G(y,n+y;d-x,e-x-y) +$$

$$(1-q_{[2]}^{y})G(0,n;d-x,y,e-x)\} +$$

$$\left(\frac{(q_{[2]}^{x}-q_1^x)(1-q_{[2]}^{d-x})}{1-q_{[2]}^{d}}\right)\{q_1^y G(x,n+x;d-x,e-x-y) +$$

$$(q_{[2]}^{y}-q_1^y)G(x,y,n+x+y;d-x,e-x-y)$$

$$+ (1-q_{[2]}^{y})G(x,n+x;d-x,y,e-x)\} +$$

$$\left(\frac{1-q_{[2]}^{x}}{1-q_{[2]}^{d}}\right)\{q_1^y G(0,n;x,e-y) + (q_{[2]}^{y}-q_1^y)G(y,n+y;x,e-y)$$

$$+ (1-q_{[2]}^{y})G(0,n;x,y,e)\}.$$

For any G-situation with $m_1 = m_2 = 0$, $d_1 = 0$, $d_2 = e \geq 2$, $n \geq 1$, we take two samples of sizes x and y either both from the non-empty defective set or one sample of size x from the non-empty defective set and one sample of size y from the conditional binomial set; hence

$$(5.1.12) \qquad G(0,n;e,e) = 1 + \min\left\{\min_{\substack{x,y\geq 1\\ x+y<e}} G^{(3)}(x,y), \quad \min_{\substack{1\leq x<e\\ 1\leq y\leq n}} G^{(5)}(x,y)\right\}$$

where $G^{(3)}(x,y)$ is given by (5.2.10) with d set equal to e and

$$(5.2.13) \qquad G^{(5)}(x,y) = \left(\frac{q^y}{1-q_{[2]}^e}\right)\Big\{q_1^x(1-q_{[2]}^{e-x})G(0,n-y;e-x,e-x)$$

$$+ (q_{[2]}^x-q_1^x)(1-q_{[2]}^{e-x})G(x,n+x-y;e-x,e-x)$$

$$+ (1-q_{[2]}^x)G(0,n-y;x,e)\Big\} + \left(\frac{1-q^y}{1-q_{[2]}^e}\right)\Big\{q_1^x(1-q_{[2]}^{e-x})G(y,n;e-x,e-x)$$

$$+ (q_{[2]}^x-q_1^x)(1-q_{[2]}^{e-x})G(x,y,n+x;e-x,e-x)$$

$$+ (1-q_{[2]}^x)G(y,n;x,e)\Big\}.$$

For any G-situation with $m_1 = m_2 = n = 0$, $d_1 = 0$, $d_2 = e \geq 3$, we take two samples of sizes x and y from the non-empty defective set; hence

$$(5.2.14) \qquad G(0,0;e,e) = 1 + \min_{\substack{x,y \geq 1 \\ x+y < e}} G^{(3)}(x,y)$$

where $G^{(3)}(x,y)$ is given by (5.2.10) with d set equal to e. For the G-situation with $m_1 = m_2 = n = 0$, $d_1 = 0$, $d_2 = e = 2$ the procedure is clear and $G(0,0;2,2) = 1$.

For any G-situation with $m_1 = m \geq 2$, $m_2 = d_1 = d_2 = 0$, $e \geq 1$, we take two samples of sizes x and y either both from the non-empty mediocre set or one sample of size x from the non-empty mediocre set and one sample of size y from the trinomial set; hence

$$(5.2.15) \qquad G(m,n;0,e) = 1 + \min\left\{\min_{\substack{x,y \geq 1 \\ x+y < m}} G^{(1)}(x,y), \quad \min_{\substack{1 \leq x < m \\ 1 \leq y \leq e}} G^{(5)}(x,y)\right\}$$

where $G^{(1)}(x,y)$ is given by (5.2.7) with d set equal to zero and

$$(5.2.16) \qquad G^{(5)}(x,y) = \left(\frac{q^x(1-q^{m-x})}{1-q^m}\right)\{q_1^yG(m-x,n-x;0,e-y)$$

$$+ (q_{[2]}^y-q_1^y)G(m-x+y,n-x+y;0,e-y)$$

$$+ (1-q_{[2]}^y)G(m-x,n-x;y,e)\}$$

$$+ \left(\frac{1-q^x}{1-q^m}\right)\{q_1^yG(x,n;0,e-y) + (q_{[2]}^y-q_1^y)G(x,y,n+y;0,e-y)$$

$$+ (1-q_{[2]}^y)G(x,n;y,e)\}.$$

Whenever we are in $G(m,n;d_1,d_2,e)$-situation with $d_1 \geqq d_2$ we proceed as in the $G(m,n;d_1,e)$-situation. For the $G(m,n;0,0)$-situation we can use the procedure for the binomial problem [5].

The boundary conditions state that for all $\vec{q} = (q_1,q_2,q_3)$

(5.2.17) $$G(1,m,n;1,d,e) = G(m,1,n;1,d,e) = G(1,m,n;d,1,e) = G(m,1,n;d,1,e)$$
$$= G(m-1,n-1;d,e-1) \text{ for } n \geqq m+1 \geqq 1,\ e \geqq d+1 \geqq 1.$$

(5.2.18) $$G(1,n;d,e) = G(0,n-1;d,e) \quad \text{for } n \geqq 1,\ e \geqq d \geqq 0.$$

(5.2.19) $$G(0,n;1,e) = H(n;e-1) \quad \text{for } n \geqq 0,\ e \geqq 1.$$

(5.2.20) $$H(0;0) = 0.$$

REFERENCES

[1] Dorfman, R. (1943). The detection of defective members of large populations. Ann. Math. Stat. 14, 436-440.

[2] Feller, W. (1957). An Introduction to Probability Theory and its Applications. (2nd edition). John Wiley and Sons, New York.

[3] Groll, P. A. and Sobel, M. (1964). Binomial group-testing with an unknown proportion of defectives. Technical Report, University of Minnesota.

[4] Huffman, David A. (1952). A method for the construction of minimum redundancy codes. Proc. I. R. E. 40, 1098-1101.

[5] Sobel, M. and Groll, P. A. (1959). Group-testing to eliminate efficiently all defectives in a binomial sample. Bell System Tech. Jour. 38, 1179-1252.

[6] Sobel, M. (1960). Group-testing to classify all defectives in a binomial sample. A contribution in Information and Decision Processes, edited by Machol, R. E., McGraw-Hill, 127-61.

[7] Sobel, M. (1964). Optimal group testing. Technical Report, Stanford University.

[8] Ungar, P. (1960). The cut-off point for group-testing. Comm. on Pure and Applied Math. 13, 49-54.

TABLE I

Values of $x_H(n;e;\vec{q}_o)$ and $x_G(m,n;d,e;\vec{q}_0)$ for $q_0 = (.90,.05,.05)$ for Procedure R_1 for various H- and G-situations arising in the classification of $N(\leq 8)$ units under R_1#.

$$x_H(n;\underset{\uparrow}{e};\vec{q}_0) = e \qquad \text{for } e \geq 1 \text{ and } n \leq N;$$

$$x_H(\underset{\uparrow}{n};0,\vec{q}_0) = n \qquad \text{for } 1 \leq n \leq N;$$

$$x_G(\underset{\uparrow}{m},n;d,e;\vec{q}_0) = \begin{cases} 1 & \text{for } m = 2,3 \\ 2 & \text{for } m = 4,5,6 \\ 3 & \text{for } m = 7 \\ 4 & \text{for } m = 8; \end{cases}$$

$$x_G(0,n;\underset{\uparrow}{d},e;\vec{q}_0) = \begin{cases} 1 & \text{for } d = 2,3 \\ 2 & \text{for } d = 4,5,6 \\ 3 & \text{for } d = 7,8. \end{cases}$$

#The vertical arrow indicates what set the x's come from.

TABLE II

Comparison of the Expected Number of Tests for Procedure R_1 and Information Theory Lower Bounds for Any Procedure starting with a Trinomial set of size N with $\vec{q}_0 = (.90,.05,.05)$.

N	H(0;N)	Information Theory Lower Bound for Any Procedure
1	1.000	0.359
2	1.288	0.718
3	1.654	1.077
4	2.034	1.436
5	2.464	1.795
6	2.905	2.154
7	3.357	2.513
8	3.820	2.872

Acknowledgement

I wish to express my sincere gratitude to Professor Milton Sobel for proposing this investigation to me and for his constant encouragement and many valuable suggestions during its progress. It has been both a privilege and a pleasure to work under his direction.

SUMMARY

The problem of group testing is concerned with the classification of N units into one of a finite number k of disjoint categories which are labeled as 'the best', 'the 2nd best', ..., 'the kth best'. A group test is a simultaneous test on x units ($1 \leq x \leq N$) with one of the following k possible outcomes: at least one of the x units belongs to the ith best category and none of the x units belongs to the jth best category for $j > i (i=1,2,\ldots,k)$. Each unit is assumed to represent an independent observation from a multinomial population with a known a priori probability q_i of any unit belonging to the ith best category for $i = 1,2,\ldots,k$ where $q_i \geq 0$ and $q_1 + q_2 + \cdots + q_k = 1$. The problem is to define a simple and efficient procedure (or an optimal procedure) for classifying all the N units by means of group tests.

For $k = 2$ the problem will be referred to as the binomial group-testing problem. The first application of binomial group-testing [1], [2] is concerned with the classification of blood for a large group of people as to whether or not each one has a particular disease. Sobel and Groll in [5] have given a procedure, which is called R_1, for the binomial group-testing problem and have investigated some properties of R_1. It is proved by Ungar [8] that for $q_1 < q^* = (\sqrt{5}-1)/2 = .618\ldots$, the units are tested one at a time under the optimal procedure; it is shown in [5] that the same property holds under R_1 for $q_1 < q^*$. In [6] Sobel has obtained the lower bounds for any group-testing procedure using information theory and the expected length of Huffman codes [4]. In [7] a procedure, called R_{-1}, is proposed and is conjectured to be optimal for all values of q_1, q_2 and N. In [3] the binomial group-testing is extended to the case in which q_1 and q_2 are unknown and a non-mixing procedure is defined and compared with other procedures.

For $k = 3$ the problem of group-testing will be referred to as the 3-category problem. In this case the three possible categories are called good, mediocre and defective.